



UNIVERSIDADE FEDERAL DE SERGIPE  
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

# **Data Analytics Transparente para Descoberta de Padrões e Anomalias na Realização de Convênios e Contratos de Repasse Federais**

Dissertação de Mestrado

Othon Stuart Ferreira Campos



São Cristóvão – Sergipe

2018

UNIVERSIDADE FEDERAL DE SERGIPE  
CENTRO DE CIÊNCIAS EXATAS E TECNOLOGIA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Othon Stuart Ferreira Campos

**Data Analytics Transparente para Descoberta de Padrões e  
Anomalias na Realização de Convênios e Contratos de  
Repasse Federais**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Sergipe como requisito parcial para a obtenção do título de mestre em Ciência da Computação.

Orientador(a): Prof. Dr. Methanias Colaço Rodrigues Júnior

São Cristóvão – Sergipe

2018

**FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL  
UNIVERSIDADE FEDERAL DE SERGIPE**

C198d Campos, Othon Stuart Ferreira  
*Data analytics* transparente para descoberta de padrões e anomalias na realização de convênios e contratos de repasses federais / Othon Stuart Ferreira Campos ; orientador Methanias Colaço Rodrigues Júnior . - São Cristóvão, 2018.

96 f. : il.

Dissertação (mestrado em Ciência da Computação) – Universidade Federal de Sergipe, 2018.

1. Computação. 2. Transparência na administração pública. 3. Governança corporativa. 4. Convênios. 5. Mineração de dados (Computação) I. Rodrigues Júnior, Methanias Colaço orient. II. Título.

CDU 004:35



UNIVERSIDADE FEDERAL DE SERGIPE  
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA  
COORDENAÇÃO DE PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Ata da Sessão Solene de Defesa da Dissertação do  
Curso de Mestrado em Ciência da Computação-UFS.  
Candidato: OTHON STUART FERREIRA CAMPOS

Em 29 dias do mês de Agosto do ano de dois mil e dezoito, com início às 14h30min, realizou-se na Sala de Reunião Superior do DCOMP da Universidade Federal de Sergipe, na Cidade Universitária Prof. José Aloísio de Campos, a Sessão Pública de Defesa de Dissertação de Mestrado do candidato **OTHON STUART FERREIRA CAMPOS**, que desenvolveu o trabalho intitulado: "**Data Analytics Transparente para Descoberta de Padrões e Anomalias na Realização de Convênios e Contratos de Repasse Federais**", sob a orientação do Prof. Dr. **Methanias Colaço Rodrigues Júnior**. A Sessão foi presidida pelo Prof. Dr. **Methanias Colaço Rodrigues Júnior** (PROCC/UFS), que após a apresentação da dissertação passou a palavra aos outros membros da Banca Examinadora, Profª. Drª. **Adicinéia Aparecida de Oliveira** (PROCC/UFS) e, em seguida, ao Prof. Dr. **Evandro de Barros Costa** (UFAL). Após as discussões, a Banca Examinadora reuniu-se e considerou o mestrando (a) APROVADO "(aprovado/reprovado)" COM "(com sem)" ressalvas. Atendidas as exigências da Instrução Normativa 01/2017/PROCC, do Regimento Interno do PROCC (Resolução 67/2014/CONEP), e da Resolução nº 25/2014/CONEP que regulamentam a Apresentação e Defesa de Dissertação, e nada mais havendo a tratar, a Banca Examinadora elaborou esta Ata que será assinada pelos seus membros e pelo mestrando.

Cidade Universitária "Prof. José Aloísio de Campos", 29 de Agosto de 2018.

Prof. Dr. Methanias Colaço Rodrigues Júnior  
(PROCC/UFS)  
Presidente

Prof. Dr. Evandro de Barros Costa  
(UFAL)  
Examinador Externo

Profª. Drª. Adicinéia Aparecida de Oliveira  
(PROCC/UFS)  
Examinador Interno

Othon Stuart Ferreira Campos  
Candidato



UNIVERSIDADE FEDERAL DE SERGIPE  
PRÓ-REITORIA DE PÓS-GRADUAÇÃO E PESQUISA  
COORDENAÇÃO DE PÓS-GRADUAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Observações (em caso de aprovação com ressalvas):

- ANÁLISE DO MEI DE DEZEMBRO E ELEIÇÕES DEVERIA IR P/ O SITE
- TRANSLAÇÕES DE OUTLIER E REGRA, COM TEXTO PADRÃO QUE MUDA SÓ O OUT E A REGRA (FILTRANDO PELA SIGNIFICÂNCIA); REVISAR PORTUGUÊS DO SOBRE DO SITE E OBSERVAR PELA BANCA NO TEXTO DAS DISCUSSÕES.
- MELHORAR JUSTIFICATIVA DA CICLOMÉTICA DE RA
- APRESENTAR LIMITAÇÕES NA CONCLUSÃO E CONTRIBUIÇÕES PARA COMPUTAÇÃO.

Prof. Dr. Methanias Colaço Rodrigues Júnior  
(PROCC/UFS)

Presidente

Prof. Dr. Evandro de Barros Costa  
(UFAL)

Examinador Externo

Profª. Drª. Adicinéia Aparecida de Oliveira  
(PROCC/UFS)

Examinador Interno

Othon Stuart Ferreira Campos

Candidato

*À minha amada esposa Anne, minha filha querida Laura e meus amorosos pais.*

# Agradecimentos

Aos meus pais, por terem me ensinado os princípios de uma verdadeira família, pelo amor incondicional e por estarem sempre ao meu lado.

Aos meus irmãos William e George, que mesmo distantes, geograficamente, mantemos a união fraterna e os laços de companheirismo.

À minha amada esposa Anne Reis pelo apoio, compreensão e paciência. Sem você ao meu lado não conseguiria concluir mais esse desafio.

À minha filhinha linda e adorada Laura, que me fortalece com os sorrisos e abraços.

Ao meu orientador, Prof. Dr. Methanias Colaço Rodrigues Júnior, por acreditar, pela disponibilidade, pelas orientações, pelos conselhos, pelas críticas e, principalmente, pelo incentivo e apoio!

À Prof<sup>a</sup> Dra. Adicinéia Aparecida de Oliveira, pelos ensinamentos compartilhados, pelas discussões em sala de aula e pelo exemplo de pessoa.

Aos amigos e companheiros de mestrado: Adriano Lima, André Santos, Erick Barros, Fabio Manguiera, Hugo Doria, Robert Oliveira e Gabriel Galvan e todos os demais que contribuíram diretamente.

E um grande obrigado a todos que, de alguma forma, contribuíram para o meu crescimento pessoal e para a realização deste trabalho.

# Resumo

**Contexto:** As recentes políticas de disponibilização de dados públicos pelo governo federal devem estimular a participação da população na fiscalização da transparência da gestão. Esta participação pode ser potencializada e estimulada, desde que os dados fornecidos pelo governo sejam estruturados e dispostos de uma forma entendível, ou efetivamente transparente, para o cidadão. Quando isto não ocorre, pesquisas sobre dados abertos podem contribuir com a estruturação, autenticação, interpretação, organização, tradução, descoberta de padrões e disponibilização de informação útil para tomada de decisão e para o controle social exercido pela população.

**Objetivo:** Este trabalho teve como objetivo aplicar técnicas de *Data Analytics* e Mineração de Dados, para analisar dados públicos sobre convênios e contratos de repasse federais, avaliando a possibilidade de criação de aplicação transparente para encontrar correlações e anomalias que evidenciem ilegalidade administrativa ([www.transparenciatraduzida.com.br](http://www.transparenciatraduzida.com.br)). **Método:** Foi realizado um mapeamento sistemático, para caracterizar os últimos 10 anos na área, um estudo de caso para identificação de *outliers*, bem como um experimento, para descoberta de regras de associação envolvendo parlamentares, com o suporte de análises estatísticas e testes de hipóteses. **Resultados:** Não foram encontrados trabalhos sobre convênios e repasses federais, nas bases científicas indexadas pelo *Scopus*. Após a execução do estudo de caso, identificou-se 36 convênios com valores aberrantes (*outliers*), totalizando R\$ 47.961.892,70. Com a execução do processo experimental, foram identificadas duas regras de associação entre parlamentares e áreas.

**Conclusão:** Constatou-se o baixo número de publicações primárias, na área de análise inteligente de dados para Governança Pública, evidenciando um grande campo para desenvolvimento de pesquisas futuras. Foram encontrados diversos valores aberrantes (*outliers*) entre os valores dos convênios ainda em execução no Brasil, inclusive no Estado de Sergipe. Em se tratando do experimento realizado para Sergipe, após aplicação do algoritmo para descoberta de regras de associação, foram encontradas apenas duas regras significantes. Estas descobertas iniciais e suas explicações, em um nível entendível para a população, evidenciam a necessidade de tradução popular dos dados abertos brutos brasileiros, legitimando o verdadeiro conceito de transparência.

**Palavras-chave:** Transparência; Governança; Dados Abertos; Convênios e Repasses Federais; *Data Analytics*.



# Abstract

**Context:** The recent political provision of public data by the Federal Government should stimulate the population participation of the management. This participation may be potentiated and stimulated, since the data given by the government are structured and understandable, or effectively transparent to the citizen. When this process does not happen, research about open data may contribute with authentication, interpretation, organization, translation, standard discovery and provision of useful information in order of social control taken by the population. **Objective:** The objective of this paper was to apply the Data Analytics and Data Mining in order to analyze public data, about agreements and federal contracts, evaluating the possibility of creating transparent applications to find correlations and abnormalities that evidence administrative illegality ([www.transparenciatraduzida.com.br](http://www.transparenciatraduzida.com.br)). **Method:** A systematic mapping was carried out to characterize the last 10 years in the field, a case study to identify outliers, as well as an experiment, to discover rules of association involving parliamentarians, supported by an analysis of results and hypotheses. **Results:** There were no papers on agreements and federal transfers, based on the scientific bases indexed by Scopus. After conducting the case study, we identified 36 agreements with aberrant values, totaling R\$ 47,961,892.70. With the accomplishment of the experimental process, were two new rules of association between parliamentarians and areas. **Conclusion:** It was verified the low number of primary publications, in an area of intelligent data analysis for Public Governance, evidencing a great field for the development of future researches. Forms of value openings (*outliers*) between the values of the agreements still in execution in Brazil, including in the State of Sergipe. The experimental testing of Sergipe, after applying algorithm for flag of the behavior of the association of flag was only a few significant rules. The findings and their explanations, in a concept perceived for the population, highlight the need for a popular translation of raw open data, legitimizing the true concept of transparency.

**Keywords:** Transparency; Governance; Open Data; Agreements and Federal Onlending; Data Analytics.

# Lista de ilustrações

Figura 1 – Distribuição de Convênios por mês. . . . .	18
Figura 2 – Percentual de convênios aprovados apenas nos meses de dezembro. . . . .	19
Figura 3 – Modelo de Dados. . . . .	22
Figura 4 – Processo de extração dos dados. . . . .	22
Figura 5 – Etapas do processo KDD. . . . .	27
Figura 6 – Gráfico Box Plot . . . . .	29
Figura 7 – Logotipo do site Transparência Traduzida. . . . .	33
Figura 8 – Distribuição dos temas abordados. . . . .	39
Figura 9 – Periódicos e Conferências. . . . .	40
Figura 10 – Distribuição dos convênios de todo o Brasil. . . . .	45
Figura 11 – Distribuição dos valores dos convênios de todo o Brasil. . . . .	45
Figura 12 – <i>Outliers</i> dos convênios de todo o Brasil. . . . .	46
Figura 13 – Dispersão dos valores dos convênios ativos no Brasil. . . . .	46
Figura 14 – Distribuição dos convênios no Estado de Sergipe. . . . .	47
Figura 15 – Distribuição de valores dos convênios no Estado de Sergipe. . . . .	47
Figura 16 – Distribuição dos valores dos convênios no Estado de Sergipe. . . . .	48
Figura 17 – <i>Outliers</i> dos convênios no Estado de Sergipe. . . . .	49
Figura 18 – <i>Outliers</i> no site Transparência Traduzida . . . . .	50
Figura 19 – Frequência das Áreas ou Parlamentares. . . . .	61
Figura 20 – Regras de associação do site: Transparência Traduzida. . . . .	62
Figura 21 – Frequência dos Parlamentares e Favorecidos. . . . .	63
Figura 22 – Frequência dos Parlamentares e Cidades. . . . .	63

# Lista de tabelas

Tabela 1 – Quantidade de Convênios e Valores Global por ano. . . . .	18
Tabela 2 – Situação de todos os convênios no período de 2008 a 2018 (R\$ bilhões) . .	19
Tabela 3 – Processo de seleção. . . . .	38
Tabela 4 – Estudos primários selecionados. . . . .	38
Tabela 5 – Técnicas específicas. . . . .	39
Tabela 6 – Periódicos e Conferências. . . . .	40
Tabela 7 – Quantidades e valores dos convênios por proponente. . . . .	51
Tabela 8 – Quantidades e valores dos convênios por área. . . . .	51
Tabela 9 – Quantidades e valores dos convênios por parlamentar. . . . .	52
Tabela 10 – Quantidades e valores dos convênios por favorecido. . . . .	52
Tabela 11 – Distribuição dos Convênios Ativos por unidade federativa (Valores em R\$ bilhões). . . . .	55
Tabela 12 – Total de Convênios e Valores dos principais Ministérios repassadores de 2008 a 2018 (R\$ bilhões) . . . . .	56
Tabela 13 – Quantitativos de transações analisadas para cada regra. . . . .	60
Tabela 14 – Regras encontradas para associações entre Parlamentares e Áreas. . . . .	61

# Lista de abreviaturas e siglas

CGU	Controladoria Geral da União
CSV	<i>Comma Separated Values</i>
e-GOV	Governo Eletrônico
ENAP	Escola Nacional de Administração Pública
ETL	<i>Extract, Transform and Load</i> - Extração, Tratamento e Carga de Dados
KDD	<i>Knowledge Discovery in Database</i>
MD	Mineração de Dados
MPOG	Ministério do Planejamento, Orçamento e Gestão
OECD	Organização para a Cooperação e Desenvolvimento Econômico
OMC	Organização Mundial de Comércio
OSCIP	Organizações da Sociedade Civil de Interesse Público
OGP	<i>Open Government Partnership</i>
SICONV	Sistema de Gestão de Convênios e Contratos de Repasse
TCU	Tribunal de Contas da União
TIC	Tecnologia da Informação e Comunicação
SEAD	Secretaria Especial de Agricultura Familiar e do Desenvolvimento Agrário
SUS	Sistema Único de Saúde

# Sumário

<b>1</b>	<b>Introdução</b>	<b>14</b>
1.1	Contextualização	14
1.2	Problemática e Hipótese	17
1.3	Justificativa	20
1.4	Objetivos da Pesquisa	20
1.4.1	Objetivos Específicos	21
1.5	Metodologia	21
1.6	Organização da Dissertação	23
<b>2</b>	<b>Fundamentação Teórica</b>	<b>24</b>
2.1	Transparência na gestão pública	24
2.2	Transferências Voluntárias e Convênios	25
2.3	Data Mining	27
2.3.1	Técnicas de Data Mining	28
2.3.1.1	Baseada em estatística	28
2.3.1.1.1	Outlier	28
2.3.1.2	Regras de Associação	30
2.4	Data Analytics	32
2.5	Transparência Traduzida	32
<b>3</b>	<b>Análise Inteligente de Dados Aplicada à Governança Pública e à Transparência de Convênios e Repasses Federais: Um Mapeamento Sistemático</b>	<b>34</b>
3.1	Método	34
3.1.1	Questões de pesquisa	35
3.1.2	Escopo da pesquisa	35
3.1.3	Método de Busca de Publicações	36
3.1.4	Critérios de Seleção	36
3.2	Ameaças à Validade	37
3.3	Discussão	37
3.4	Considerações Finais do Capítulo	41
<b>4</b>	<b>Estudo de caso de descrição dos Dados de Convênios Brasileiros</b>	<b>42</b>
4.1	Definição do objeto	42
4.2	Planejamento	42
4.2.1	Seleção de contexto	43
4.2.2	Instrumentação	43

4.3	Preparação . . . . .	43
4.4	Execução . . . . .	44
4.4.1	Coleta de dados . . . . .	44
4.4.2	Detecção de Outliers . . . . .	44
4.5	Análise e interpretação . . . . .	44
4.6	Considerações finais do capítulo . . . . .	53
<b>5</b>	<b>Experimento sobre Associações . . . . .</b>	<b>54</b>
5.1	Definição e planejamento do experimento . . . . .	54
5.1.1	Definição do Objetivo . . . . .	54
5.1.2	Planejamento . . . . .	57
5.2	Operação do Experimento . . . . .	58
5.2.1	Preparação . . . . .	58
5.2.2	Execução . . . . .	59
5.2.3	Coleta de Dados . . . . .	59
5.2.4	Validação dos Dados . . . . .	59
5.3	Resultados . . . . .	59
5.3.1	Análise e Interpretação . . . . .	59
5.3.1.1	Associação entre Parlamentares e Áreas . . . . .	60
5.3.1.2	Associação entre Parlamentares e Favorecidos . . . . .	62
5.3.1.3	Associação entre Parlamentares e Cidades . . . . .	62
5.3.2	Ameaças à Validade . . . . .	64
5.3.2.1	Ameaças à validade interna . . . . .	64
5.3.2.2	Ameaças à validade externa . . . . .	64
5.4	Considerações Finais . . . . .	64
<b>6</b>	<b>Conclusão . . . . .</b>	<b>65</b>
6.1	Contribuições . . . . .	65
6.2	Limitações . . . . .	66
6.3	Perspectivas . . . . .	67
6.4	Considerações Finais . . . . .	67
	<b>Referências . . . . .</b>	<b>68</b>
	 <b>Apêndices</b>	 <b>71</b>
	<b>APÊNDICE A Script: Criação do base de dados . . . . .</b>	<b>72</b>

APÊNDICE B	Comandos SQL para listagem dos dados dos convênios classifica- dos como <i>outliers</i> e comando para geração do arquivo CSV utili- zado no gráfico de <i>outliers</i> . . . . .	74
APÊNDICE C	Comandos SQL para geração dos arquivos com os conjuntos fre- quentes . . . . .	76
APÊNDICE D	Comandos do R para estudo dos outliers . . . . .	78
APÊNDICE E	Comandos do R para estudo das Regras de Associação . . . . .	79
APÊNDICE F	Scripts em Pythons . . . . .	81
<b>Anexos</b>		<b>86</b>
ANEXO A	Dicionário de dados . . . . .	87
ANEXO B	Lista dos <i>outliers</i> . . . . .	91

# 1

## Introdução

Este capítulo realiza uma breve contextualização relacionada ao tema da pesquisa, motivação, problemática, questões, objetivos, hipótese e suposição que se pretende evidenciar. Além disso, são descritas as contribuições que se espera alcançar ao final do trabalho e a metodologia de pesquisa direcionadora do trabalho. Por fim, apresenta-se a organização da dissertação.

### 1.1 Contextualização

Em um Governo Democrático, os princípios políticos são baseados na soberania popular e da distribuição equitativa do poder. Na Democracia Representativa o povo tem o poder e a classe política, representantes eleitos pelo povo, possui um mandato político temporário. Neste regime político, uma das premissas é que o governo deva ser visível tendo seus atos e todas as informações referentes às suas administrações tornados públicos, para fortalecimento da transparência perante aos seus representados ([URBINATI, 2006](#)).

Atualmente, as discussões sobre a ineficiência do estado ante os gastos públicos e corrupção tem recebido mais atenção da sociedade civil e tem aparecido com maior recorrência nas pautas políticas. A ineficiência nos gastos públicos impacta diretamente nas políticas públicas e no crescimento do país, tornando o tema Transparência mais relevante e a transformando em uma ferramenta de controle dos governantes ([MATHEUS; JANSSEN, 2016](#)).

Concomitante ao incremento na demanda por maior transparência houve uma substancial evolução na Tecnologia da Informação e Comunicação (TIC) o que ajudou a criar novas maneiras de disponibilizar informações públicas à população. Novos sistemas foram desenvolvidos, novos serviços oferecidos e integrações sistêmicas aconteceram. Todas as revoluções tecnológicas juntamente com a popularização da internet geraram mudanças nos processos internos e nas relações do governo com o público externo. Tais mudanças podem ser chamadas de Governo



Eletrônico, ou simplesmente e-Gov.

Com o e-Gov, os órgãos públicos estão entre os maiores criadores e coletores de dados em muitos domínios. Estes domínios de dados variam de tráfego, clima, informações geográficas, turísticas, estatísticas, negócios, orçamentação do setor público, entre outros (JANSSEN; CHARALABIDIS; ZUIDERWIJK, 2012). No contexto público, o acesso à informação permite criar uma sociedade democrática com os cidadãos participativos, proporcionando-lhes ferramentas para compreender e utilizar a informação disponível e começar a cunhar um pensamento crítico sobre informações e serviços (HARRISON et al., 2012).

Ao redor do mundo vários atos legislativos foram produzidos para criar mais transparência. No setor de comércio exterior, a Organização Mundial de Comércio (OMC) firmou compromisso, em 1994, pela melhoria na transparência por meio da sua comunicação interna e externa. A Organização para a Cooperação e Desenvolvimento Econômico (OECD), em 2009, aprovou a convenção que visa combater o suborno de funcionários públicos estrangeiros nas transações comerciais internacionais (TURNES; ERNST, 2015). No setor financeiro pode-se citar o acordo de Basiléia firmado entre várias nações do mundo com o objetivo de definir mecanismos para mensurar o risco em operações de crédito estabelecendo exigências mínimas de capital para suportá-lo. Atualmente, o acordo é apoiado em 25 princípios de supervisão bancária, buscando mais transparência entre as instituições financeiras (SETTLEMENTS, 2006).

Há ainda iniciativas internacionais como a *Open Government Partnership* (OGP) criada com a finalidade de promover transparência, aumento da participação cívica, combater a corrupção e aproveitar novas tecnologias para tornar governos de diversos países mais abertos, eficazes e responsáveis (FREITAS; DACORSO, 2014). Outro exemplo é a iniciativa *Transparency International* (INTERNATIONAL, 2016) e a Transparência Brasil (BRASIL, 2016b), criadas para combater a corrupção por meio de redes de discussão sobre transparência e divulgação de ferramentas para monitorar a corrupção.

No cenário do governo federal algumas medidas foram tomadas na tentativa de evidenciar transparência nas ações públicas. Em 2001 foi criada a Corregedoria Geral da União (CGU), através da medida provisória Nº 2.143-31 (BRASIL. Palácio do Planalto, 2001), que tinha o propósito de evitar, no âmbito do Poder Executivo Federal, fraudes, corrupção e promover a defesa de bens públicos. Em 2003, com a publicação da lei Nº 10.683, a CGU ganhou status de ministério para o controle e a transparência, passando a chamar-se de Controladoria Geral da União (BRASIL. Palácio do Planalto, 2003). O decreto Nº 5.683 (BRASIL. Palácio do Planalto, 2006), deu a CGU à autoridade na detecção da corrupção e a possibilidade de antecipar-se a esta, desenvolvendo meios para prevenir a sua ocorrência. Com a publicação do decreto Nº 8109 (BRASIL. Palácio do Planalto, 2013) novas mudanças foram realizadas na CGU, com a intenção de dispor de uma estrutura compatível com a dimensão e relevância de suas competências, em especial, com as leis de Acesso à Informação, de Conflito de Interesses e Anticorrupção (MATHEUS; JANSSEN, 2016). Atualmente, após a publicação do decreto Nº 8.910 (BRASIL.

[Palácio do Planalto, 2016](#)), a CGU passou a ser o Ministério da Transparência, Fiscalização e Controladoria-Geral da União (CGU).

Ainda no âmbito do governo brasileiro podemos citar algumas outras leis criadas para promover transparência: a Lei de Acesso à Informação 12.527/2011, que regulamenta o direito constitucional de acesso às informações públicas; a Lei complementar 131/2009, que determina a disponibilização de informações sobre a execução orçamentária e financeira das instituições públicas; o Decreto 6.932/2009, que estabelece a Carta de Serviços ao Cidadão e apresenta informações sobre os serviços prestados pelos órgãos públicos.

Outra iniciativa do governo federal, que promove transparência, é o Portal dos Convênios lançado em 2008. Surgiu a partir de uma determinação do Tribunal de Contas da União (TCU) ao Ministério do Planejamento, Orçamento e Gestão (MPOG), que apresentasse um estudo técnico para implementação de sistema de informática em plataforma web, o qual permitisse o acompanhamento *on-line* de todos os convênios e outros instrumentos jurídicos utilizados para transferir recursos federais a outros órgãos/entidades, entes federados e entidades do setor privado, que pudesse ser acessado por qualquer cidadão, por meio da internet, contendo informações relativas aos instrumentos celebrados <sup>1</sup>. O Portal dos Convênios é o sítio eletrônico que abriga o Sistema de Gestão de Convênios e Contratos de Repasse (SICONV), bem como todo o conjunto de informações relacionadas aos Convênios e Contratos de Repasse tais como: Legislação (Leis, Decretos, Portarias, Instruções Normativas e Jurisprudência); Manuais, Glossário, Perguntas Frequentes, Tutoriais e informações sobre treinamentos no SICONV; Comunicados, Informações Gerenciais, Atas e Diretrizes da Comissão Gestora do SICONV. O Portal dos Convênios promove o encontro entre todos os atores envolvidos no processo de transferências voluntárias, dinamizando essa relação e oferecendo à sociedade transparência e controle social sobre a aplicação dos recursos públicos ([SICONV, 2016](#)).

Entretanto, nenhuma das iniciativas reguladoras ou sistemas que promovam a transparência parece ser suficiente para tornar uma organização transparente ao grande público. São milhares de fluxo de dados publicados diariamente sem a certeza de que sejam utilizáveis ou com a clara percepção do que está sendo divulgado. Muitas vezes sem métricas que possibilitem ao cidadão inferir conclusões acerca da publicação.

Várias estratégias são utilizadas visando a prevenção e o combate à corrupção. No entanto, devido ao crescente aumento de informações nos bancos de dados governamentais, a tarefa de exploração desses dados para geração de conhecimento útil na atividade de auditoria se torna cada vez mais árdua. As técnicas de mineração de dados têm sido alvo de várias pesquisas motivadas pelos resultados no processo de descoberta de conhecimento em grandes volumes de dados ([SILVA; ROCHA, 2010](#)).

Em razão disso, a proposta deste trabalho é conduzir um processo experimental, seguindo

---

<sup>1</sup> acessar o site (<http://www.enap.gov.br>) para maiores detalhes

as diretrizes de [Wohlin et al. \(2012\)](#) para realizar uma análise extrínseca dos dados públicos sobre projetos, lançados no portal de convênios do SICONV, avaliando a possibilidade de criação de aplicação transparente que infira padrões e identifique se há indícios de problemas nesses projetos. Para apoiar a realização desta pesquisa serão utilizadas técnicas de *Data Mining* e de regras de associação para identificação, em dados, de padrões válidos e potencialmente úteis.

## 1.2 Problemática e Hipótese

O artigo 241 da Carta Magna regula de forma explícita o conceito de cooperação associativa e disciplina a competência comum entre a União, os Estados, Distrito Federal e os Municípios:

Art. 241. A União, os Estados, o Distrito Federal e os Municípios disciplinarão por meio de lei os consórcios públicos e os convênios de cooperação entre os entes federados, autorizando a gestão associada de serviços públicos, bem como a transferência total ou parcial de encargos, serviços, pessoal e bens essenciais à continuidade dos serviços transferidos ([BRASIL, 1988](#)).

Segundo [OLIVEIRA et al. \(2011\)](#), a partir da nova Constituição Federal houve grande demanda social por obras e serviços e o Estado não supriu suficientemente essa demanda, desse modo, o repasse de recursos via convênios com estruturas de governos subnacionais e entidades privadas sem fins lucrativos, tornou-se instrumento preferencial acompanhando a intensa municipalização e a nova distribuição de competências dos entes federativos derivada da nova Constituição Federal. O SICONV surgiu exatamente para dar celeridade aos repasses via Convênios, bem como instrumento de transparência das atividades realizadas nas transações feitas pela União. Todos os atos de credenciamento, celebração, alteração, liberação de recursos, acompanhamento da execução de projetos e prestação de contas somente serão viabilizados no Portal, via internet. Ao automatizar as transferências voluntárias, o uso do SICONV confere mais transparência aos repasses de recursos públicos, uma vez que a sociedade poderá acompanhar a execução dos contratos pela internet e verificar a aplicação do montante repassado a um determinado ente da federação.

A tabela 1 mostra a distribuição do número de convênios, assinados e em execução, e o valor global de recursos financeiros considerando o período de 2008 a 2018, totalizando aproximadamente 64 mil convênios com os respectivos valores na ordem de R\$ 64 bilhões, envolvendo convênios com estados, municípios e entidades sem fins lucrativos. Nota-se que o ano de 2013, apesar de não possuir o maior número de convênios assinados, concentrou aproximadamente 21,6% dos recursos disponibilizados.

Tabela 1 – Quantidade de Convênios e Valores Global por ano.

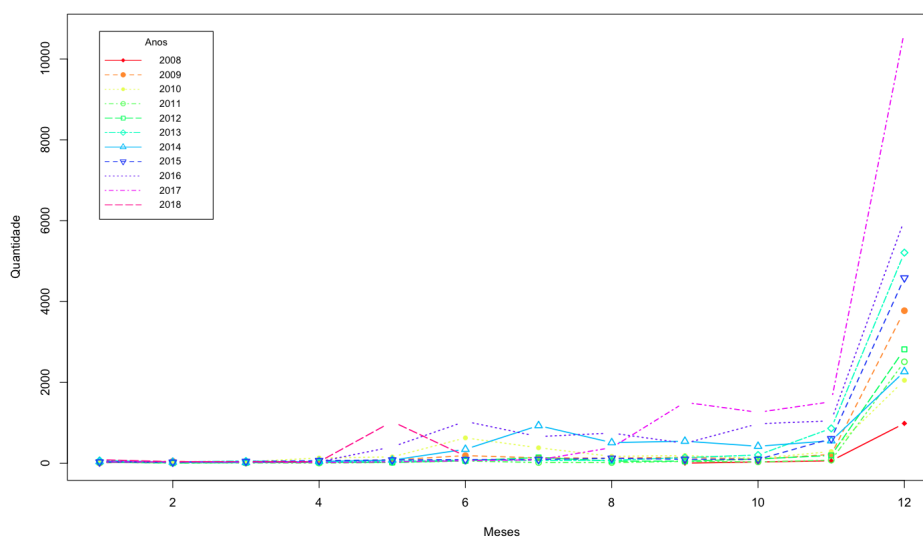
Ano	Quant.	%	Valor Total	%
2008	1083	1,7%	R\$ 1.754.563.342,99	2,5%
2009	4897	7,7%	R\$ 5.724.702.613,77	8,3%
2010	4216	6,6%	R\$ 4.975.604.153,11	7,2%
2011	2823	4,4%	R\$ 5.178.097.951,75	7,5%
2012	3516	5,5%	R\$ 7.550.589.335,48	10,9%
2013	6888	10,8%	R\$ 14.902.985.035,01	21,6%
2014	5794	9,1%	R\$ 4.787.487.904,28	6,9%
2015	5941	9,3%	R\$ 4.942.517.481,57	7,2%
2016	11544	18,1%	R\$ 7.267.109.646,71	10,5%
2017	15655	24,5%	R\$ 10.629.193.292,91	15,4%
2018	1414	2,2%	R\$ 1.331.650.791,67	1,9%
Total	63771	100%	R\$ 69.044.501.549,25	100%

Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

A figura 1 mostra a distribuição do número de convênios ao longo dos meses no período de 2008 a 2018. Observa-se uma linearidade nas quantidades até o quinto mês de cada ano, a partir deste mês um leve incremento acontece até o mês de dezembro quando em todos os anos há um aumento considerável na aprovação de convênios. Excetuando os anos de 2010 e 2014, todos os outros anos apenas o mês de dezembro é responsável por mais de 50% dos convênios. Como pode ser observado na figura 2.

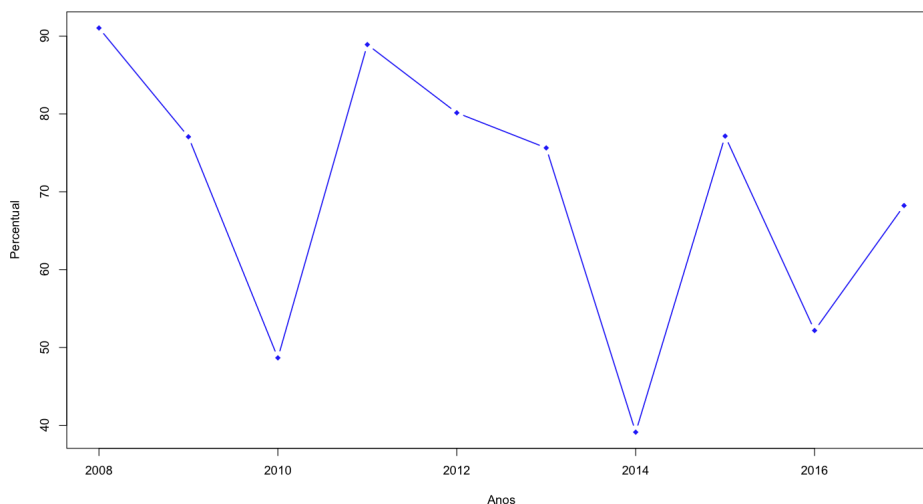
Figura 1 – Distribuição de Convênios por mês.



Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

Figura 2 – Percentual de convênios aprovados apenas nos meses de dezembro.



Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

Do total de convênios celebrados neste período, 43.472, ou 27,1%, ainda estão em execução. 13.179, ou 8,2%, foram anulados, e 27.862, o que representa aproximadamente 17,4%, dos convênios, apresentam alguma irregularidade na prestação de contas, conforme demonstrado na tabela 2.

Tabela 2 – Situação de todos os convênios no período de 2008 a 2018 (R\$ bilhões)

Situação	Quant.	%	Valor	%
Em execução	43472	27,1%	48,2	39,5%
Convênio Anulado	13179	8,2%	9,1	7,5%
Cancelado	4874	3,0%	4,9	4,0%
<b>Problemas na Prestação de Contas</b>	<b>27862</b>	<b>23,8%</b>	<b>25,50</b>	<b>34,5%</b>
<i>Aguardando Prestação de Contas</i>	5664	3,5%	4,83	4,0%
<i>Prestação de Contas enviada para Análise</i>	7309	4,6%	7,61	6,2%
<i>Prestação de Contas Aprovada com Ressalvas</i>	6696	4,2%	3,83	3,1%
<i>Prestação de Contas em Análise</i>	4365	2,7%	5,52	4,5%
<i>Prestação de Contas em Complementação</i>	1793	1,1%	2,06	1,7%
<i>Inadimplente</i>	575	0,4%	0,42	0,3%
<i>Prestação de Contas Rejeitada</i>	1460	0,9%	1,19	1,0%

Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

Baseada nesses dados, a presente proposta tem em seu escopo a análise do rol de informações disponíveis nos bancos de dados sobre convênios federais (SICONV). Neste contexto, alguns questionamentos que rodeiam esta análise são:

- Q1: Existem discrepâncias entre os valores dos convênios firmados?
- Q2: Há associações de parlamentares com áreas específicas?
- Q3: Há associações entre parlamentares e favorecidos?
- Q4: Há associações entre parlamentares e cidades específicas?

A partir dessas questões gerais, algumas hipóteses serão testadas para o estado de Sergipe:

- **Hipótese 1 (Q2)**
  - H0: Os parlamentares são independentes das áreas específicas.
  - H1: Os parlamentares são dependentes de áreas específicas.
- **Hipótese 2 (Q3)**
  - H0: Os parlamentares são independentes de empresas específicas.
  - H1: Os parlamentares são dependentes de empresas específicas.
- **Hipótese 3(Q4)**
  - H0: Os parlamentares são independentes de cidades específicas.
  - H1: Os parlamentares são dependentes de cidades específicas.

## 1.3 Justificativa

Este projeto busca realizar uma avaliação experimental que permita mensurar os dados abertos obtidos no portal de convênios (SICONV, 2016), buscando discrepâncias e anomalias nos convênios. Assim, trabalhos posteriores poderão propor melhorias nas escolhas de projetos e incrementar a apresentação das informações abertas, tornando-as mais simples e acessíveis ao grande público.

## 1.4 Objetivos da Pesquisa

O objetivo deste projeto é aplicar técnicas de *Data Analytics* e Mineração de Dados, para analisar dados públicos sobre convênios e contratos de repasse federais, avaliando a possibilidade de criação de aplicação transparente para encontrar correlações e anomalias que evidenciem ilegalidade administrativa.

### 1.4.1 Objetivos Específicos

Os objetivos norteadores do presente trabalho são os expostos abaixo:

- Desenvolver um módulo ETL (*Extract, Transform, Load*) para carregar um Data Mart sobre o SICONV;
- Mapeamento Sistemático sobre *Data Analytics*, *Big Data*, Transparência, Transferências e Repasses Públicos;
- Realizar de um estudo de caso para investigar a existência de valores aberrantes entre os convênios firmados;
- Realizar de experimento para investigar a existência de correlações e anomalias que evidenciem ilegalidade administrativa;
- Desenvolver aplicação pública que torne mais fácil a consulta e visualização dos resultados.

## 1.5 Metodologia

A metodologia adotada para o trabalho envolveu, inicialmente, uma revisão da literatura com abordagens sistemáticas, tendo por finalidade encontrar o estado da arte das pesquisas sobre Transparência Pública, Transferências e Repasses Públicos, Convênios, Governança, *Data Mining*, *Data Analytics*, *Outliers*, *Big Data* e Regras de Associação.

Para operacionalizar a revisão, acessamos a base Scopus por meio do portal de periódicos da CAPES <sup>2</sup>, pois ele permite fazer download dos artigos sem restrições. A Scopus foi escolhida por incluir buscas em diferentes bancos de dados científicos (IEEE, ACM e outros). Ato contínuo será realizado um trabalho de automação da extração, transformação e carga (ETL) dos dados brutos disponibilizados no site do SICONV <sup>3</sup>. Normalmente, essas informações são fornecidas em arquivos muito grandes, com dados dispersos. Isso limita o entendimento do cidadão com relação ao significado ou à importância dos dados abertos. O trabalho permitirá fazer a transição dos dados brutos do governo para informações estruturadas (DESSI et al., 2016).

O tratamento destes dados se iniciará com o *download* dos arquivos, no formato CSV. Cada arquivo será lido, interpretado e seu conteúdo salvo numa base de dados estruturada, ver figura 3, refletindo o modelo disponibilizado pelo próprio portal de convênios.

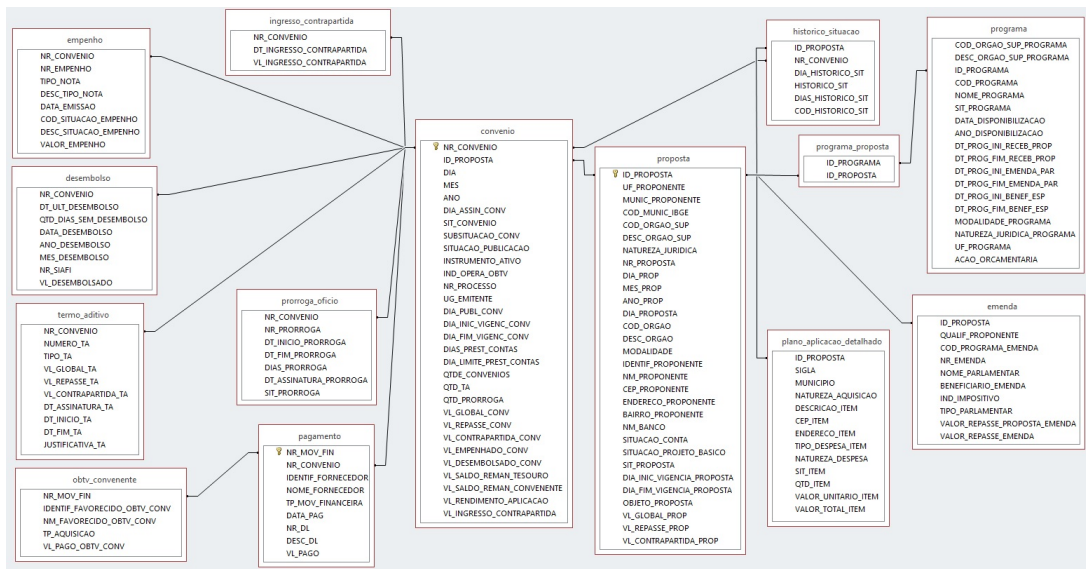
A figura 4 mostra a estrutura metodológica para construção da base estruturada de dados a partir dos arquivos brutos disponibilizados pelo portal de convênios.

<sup>2</sup> <https://www.periodicos.capes.gov.br>

<sup>3</sup> <http://portal.convenios.gov.br>



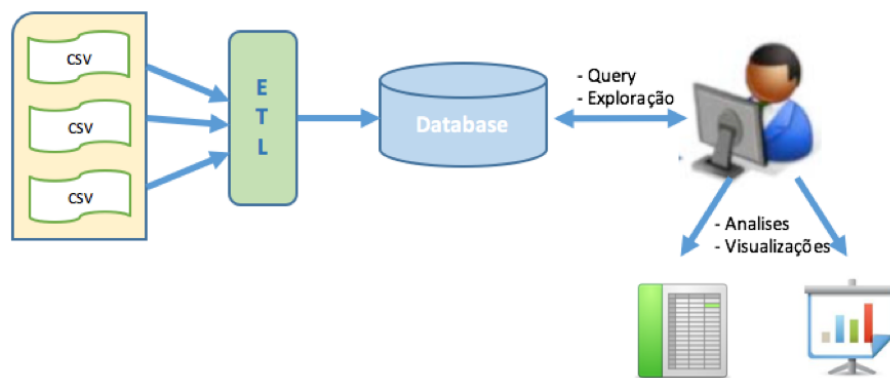
Figura 3 – Modelo de Dados.



Fonte: Portal de Convênios. (2018).

Nota: Disponível em: <http://portal.convencios.gov.br/download-de-dados>.

Figura 4 – Processo de extração dos dados.



Fonte: Elaborada pelo autor (2017).

Na sequência, será executado um estudo de caso para pesquisar discrepâncias e anomalias nos convênios firmados.

Por fim, será realizado um experimento controlado com os dados estruturados, para produção da pesquisa e coleta de dados dependentes. De acordo com Wohlin et al. (2012) uma experimentação não é uma tarefa simples, pois envolve preparar, conduzir e analisar experimentos corretamente. Os autores destacam como uma das principais vantagens da experimentação o controle dos sujeitos, objetos e instrumentação, o que torna possível extrair conclusões mais gerais sobre o assunto investigado.



No contexto experimental adotado por esta proposta, outras vantagens incluem a habilidade de realizar análises estatísticas, utilizando métodos de teste de hipóteses e oportunidades para replicação. Além disso, [Juristo e Moreno \(2013\)](#) também afirmam que a pesquisa científica, como a desta pesquisa em tela, não pode ser baseada em opiniões ou interesses comerciais. Investigações científicas são representadas por estudos baseados em observação e/ou experimentação acerca do mundo real e seus comportamentos mensuráveis.

Sumarizando, o experimento terá 3 etapas: (1) identificação e *download* dos arquivos no formato CSV; (2) construção de ETL e da base de dados para tratamento, estruturação e armazenamento das informações contidas nos arquivos CSV; (3) planejamento, instrumentação, operação e exploração dos dados estruturados.

## 1.6 Organização da Dissertação

Este documento está organizado em 6 capítulos que fornecem uma base conceitual e experimental para o seu entendimento. Os tópicos a seguir descrevem o conteúdo de cada um dos capítulos:

- O Capítulo 1 apresenta esta Introdução, explicando as justificativas juntamente com as hipóteses levantadas;
- O capítulo 2 apresenta a fundamentação teórica que fornece os conceitos de Transparência Pública, Convênios, Transferências voluntárias, e detalha e define *Outlier* e Regras de Associação;
- O capítulo 3 descreve o Mapeamento Sistemático utilizado para identificar e sistematizar as abordagens e técnicas computacionais de análise inteligente de dados que têm sido utilizadas para dar transparência, especificamente, aos repasses federais ou para favorecer a Governança Pública em geral;
- O capítulo 4 mostra um estudo de caso para identificar valores aberrantes nos dados dos convênios ainda em execução;
- O capítulo 5 apresenta o processo experimental conduzido para avaliar associações entre os convênios e os parlamentares;
- O capítulo 6, por fim, revisa todo o trabalho desenvolvido, discute as ideias centrais apresentadas da dissertação, apresenta as limitações e contribuições do trabalho, propõe linhas de trabalhos futuros e apresenta os resultados obtidos.

# 2

## Fundamentação Teórica

Este capítulo efetua uma trajetória pelos principais conceitos acerca do tema proposto, procurando embasamento para sustentar o trabalho. Aqui são definidos os termos Transparência, Convênios, Regras de Associação e *Outliers*.

### 2.1 Transparência na gestão pública

Transparência é um conceito de divulgação de informação utilizado em diferentes contextos, principalmente relacionados à capacitação dos cidadãos em relação aos seus direitos. Além disso, a transparência organizacional estabelece um conjunto de aspectos que sugerem a existência de políticas, procedimentos e tecnologias que proporcionem acesso, uso, qualidade, compreensão e auditabilidade de processos e informações ([ALBUQUERQUE et al., 2016](#)).

A transparência é um dos princípios da governança pública e as iniciativas que visem aperfeiçoar os mecanismos de transparência de informações acerca da gestão são consideradas boas práticas de governança. De modo geral, deve caracterizar todas as atividades realizadas pelos gestores públicos de maneira que os cidadãos tenham acesso e entendimento das realizações dos gestores governamentais. Podem-se destacar algumas características acerca da Transparência: informação livre, disponível, compreensível, diretamente acessível aos que serão afetados pelas decisões delas decorrentes, prestadas de forma completa em meios de comunicação adequados. Além disso, a palavra “transparência” traz consigo um importante conjunto de associações morais e políticas, entre as quais: honestidade, lisura e abertura ([CRUZ et al., 2012](#)).

Ainda sobre o conceito de transparência, para sua prática, é necessário o aprofundamento de sua semântica, da existência de métodos que possibilitem estabelecê-la de forma eficiente e sistemática nas organizações e de instrumentos que permitam verificar que a forma como é estabelecida em determinada organização atende tanto a estratégia da organização como as necessidades da sociedade ([HARRISON et al., 2012](#)). A transparência é vista como a superação

da assimetria de informação entre o governo e o público, o que deve resultar na capacidade do público de entender os vários aspectos do governo e seu funcionamento interno (MATHEUS; JANSSEN, 2016).

## 2.2 Transferências Voluntárias e Convênios

Entende-se por transferências voluntárias os recursos financeiros repassados pela União a Estados, Distrito Federal, municípios e entidades privadas sem fins lucrativos, em decorrência da celebração de convênios, contratos de repasse, termos de parceria, termos de colaboração e termos de fomento, para a realização de obras e/ou serviços de interesse comum e que não se origine de determinação constitucional ou legal, ou destine-se ao Sistema Único de Saúde (SUS) (BRASIL, 2016a).

- Convênio: É todo e qualquer instrumento formal que discipline a transferência de recursos financeiros dos orçamentos da União para um órgão ou entidade da administração pública estadual, distrital ou municipal, direta ou indireta, ou ainda entidade filantrópica sem fins lucrativos na área da saúde (§1º do art. 199 da CF/1988). Sua finalidade é a execução de programa de governo envolvendo a realização de projeto, atividade, serviço, aquisição de bens ou evento de interesse recíproco, em regime de mútua cooperação (BRASIL, 2016a).
- Contrato de repasse: Instrumento administrativo usado na transferência dos recursos financeiros por intermédio de instituição ou agente financeiro pública federal, que atua como mandatário da União. A instituição que mais fortemente vem operando essa modalidade de transferência é a Caixa Econômica Federal ([www.caixa.gov.br](http://www.caixa.gov.br)) (BRASIL, 2016a).
- Termo de parceria: Instrumento jurídico previsto na Lei 9.790/1999 para transferência de recursos para Organizações da Sociedade Civil de Interesse Público (OSCIP) (BRASIL, 2016a).

Todas as parcerias estabelecidas pela administração pública com organizações da sociedade civil para a consecução de finalidades de interesse público e recíproco que não envolvam a transferência de recursos financeiros são denominadas de **Acordo de cooperação**.

Os partícipes das transações de transferências voluntárias são definidos como:

- Proponente: órgão ou entidade pública ou privada sem fins lucrativos que manifeste, por meio de proposta de trabalho, interesse em firmar convênio, contrato de repasse, termo de parceria, termo de colaboração ou termo de fomento.
- Concedente: órgão ou entidade da administração pública federal, direta ou indireta, responsável pela transferência dos recursos financeiros ou pela descentralização dos créditos orçamentários destinados à execução do objeto do convênio.

- **Conveniente:** órgão ou entidade da administração pública, direta ou indireta, de qualquer esfera de governo, consórcio público ou entidade privada sem fins lucrativos, com o qual a administração federal pactua a execução de programa, projeto, atividade ou evento mediante convênio.
- **Contratante:** órgão ou entidade da administração pública federal, direta ou indireta, que pactua a execução de programa, projeto, atividade ou evento, por intermédio de instituição financeira federal (mandatária) mediante celebração de contrato de repasse.
- **Contratado:** órgão ou entidade da administração pública, direta ou indireta, de qualquer esfera de governo, consórcio público ou entidade privada sem fins lucrativos, com a qual a administração federal pactua a execução de contrato de repasse.

Os recursos para execução dos convênios têm origem de duas maneiras:

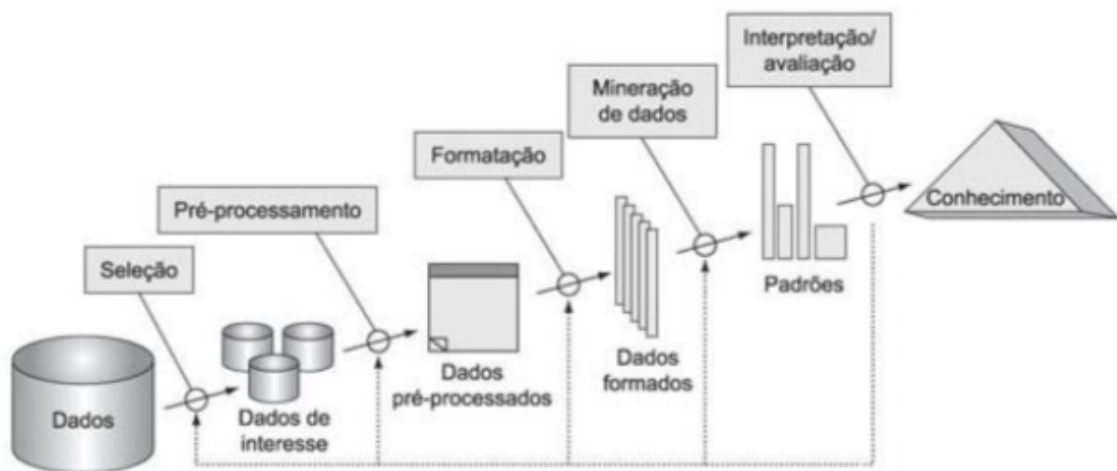
- **Contemplação nominal do Estado, do Município ou da entidade privada sem fins lucrativos** por meio da proposta do Executivo ou de emenda ao Orçamento da União por Deputado Federal ou Senador.
- **Não contemplação explícita**, mas o programa orçamentário destina recursos para a região onde se localiza o pretendente e prevê sua aplicação por meio de órgão ou entidade estadual, municipal ou não governamental.

A emenda parlamentar pode ser proposta por um parlamentar (Deputado Federal ou Senador), uma Bancada (Termo informal para representar um agrupamento organizado dos parlamentares de uma mesma representação partidária) ou ainda uma Comissão (As comissões são órgãos integrados por deputados, com composição partidária proporcional à da Câmara, que podem ter caráter permanente ou temporário). Para esse estudo foram selecionadas apenas as emendas individuais, assim a pesquisa pode focar na figura única do parlamentar.

## 2.3 Data Mining

Segundo (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996), a mineração de dados refere-se a um passo particular no processo de KDD (*Kn Discovery in Database*) que consiste na aplicação de análise de dados e algoritmos específicos para extração de padrões de dados. O KDD refere-se ao processo geral de descoberta de conhecimento útil a partir de dados, como demonstrado da figura 5.

Figura 5 – Etapas do processo KDD.



Fonte: Baseado em Fayyad, Piatetsky-Shapiro e Smyth (1996)

O processo torna-se iterativo na medida em que os resultados obtidos por meio da mineração de dados fazem pouco ou nenhum sentido, exigindo assim um recalibramento das funções de mineração relacionando a parte interativa do processo pelo qual se obtém dados que sejam de fato úteis ao negócio (SILVA; ROCHA, 2010).

Ainda segundo (SILVA; RALHA, 2011), a utilização da técnica de mineração de dados (MD) tem se mostrado de grande valia na obtenção de informações e no processo de descoberta de conhecimento em cima de grandes volumes de dados. Consiste basicamente na especificação de métodos que nos garantam como descobrir os padrões que nos interessam. As tarefas de MD são geralmente divididas em duas categorias principais, a de tarefas preditivas e a de tarefas descritivas. A primeira objetiva prever o valor de um atributo particular baseado nos valores de outros atributos. O atributo a ser predito é conhecido como alvo ou variável dependente, enquanto os atributos usados para fazer a predição são conhecidos como explanatórios ou variáveis independentes. Já a segunda categoria tem como objetivo derivar padrões como correlações, tendências, grupos, trajetórias e anomalias as quais sumarizam as relações subjacentes nos dados.

### 2.3.1 Técnicas de Data Mining

Existem diversas técnicas de Mineração de Dados que são usadas para encontrar os tipos de padrões a serem encontrados em um conjunto de dados. As principais técnicas e os tipos de padrões que elas podem descobrir são descritos abaixo.

#### 2.3.1.1 Baseada em estatística

##### 2.3.1.1.1 Outlier

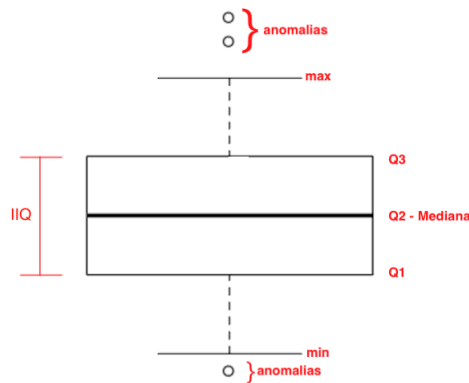
Os *outliers* são algo fora do comum e que muitas vezes pode representar anomalias que evidenciem ilegalidades administrativas. Segundo [Conforti, Rosa e Hofstede \(2015\)](#) o conceito do *Outlier* é usado para a rotulagem e identificação das observações que se afastem demasiadamente de outras observações. Baseia-se em modelos estatísticos, dos quais se espera que os dados estejam em uma distribuição normal e as observações improváveis, com base na média e desvio padrão, são consideradas como valores aberrantes ou anormais.

Um conjunto de dados pode conter objetos que não estão em conformidade com o comportamento geral ou modelo dos dados. Esses objetos de dados são *outliers*. Muitos métodos de mineração de dados descartam *outliers* como ruído ou exceções. No entanto, em algumas aplicações (por exemplo, detecção de fraude), os eventos raros podem ser mais interessantes do que os mais frequentes. A análise de dados discrepantes é referida como análise *outlier* ou mineração de anomalias. Tais exceções podem ser detectadas usando testes estatísticos que pressupõem um modelo de distribuição ou probabilidade para os dados ou usando medidas de distância em que os objetos que são remotos de qualquer outro *cluster* são considerados *outliers*. Em vez de usar medidas estatísticas ou de distância, os métodos baseados em densidade podem identificar *outliers* em uma região local, embora pareçam normais a partir de uma visão de distribuição estatística global ([HAN; PEI; KAMBER, 2011](#)).

*BoxPlot* é, provavelmente, a técnica estatística mais simples para se detectar *outliers* e representá-los graficamente. A construção desta representação gráfica utiliza a sumarização de 5 números: a menor representação não anômala (**mim**), o primeiro quartil (**Q1**), a linha central (**mediana** ou **Q2**), o quartil superior (**Q3**), a maior representação não anômala (**max**) e, por fim, o intervalo interquartil (**IIQ**) que representa a diferença entre **Q3** e **Q1**. Para se determinar os valores dos limites inferior e superior, representados por (**mim**) e (**max**) respectivamente, ([LAURIKKALA et al., 2000](#)) sugerem a utilização de uma heurística ( $1.5 \times \text{IIQ}$ ). Assim o limite inferior é calculado por  $\text{Q1} - (1.5 \times \text{IQR})$  e o limite superior é dados por  $\text{Q3} + (1.5 \times \text{IQR})$ . As anomalias são representadas pelos pontos encontrados além dos valores **mim** e **max**. A figura 6 mostra um exemplo de *BoxPlot* e seus valores. Os *boxplots* caracterizam uma amostra usando os percentis 25%, 50% e 75%, associados aos quartis (**Q1**), (**Q2**) e (**Q3**) respectivamente.

O valor 1.5, utilizado para o cálculo do intervalo de confiança (**mim** e **max**), é utilizado para representar 95% dos dados e os extremos razoáveis. Utilizando o valor 3.0 conseguimos

Figura 6 – Gráfico Box Plot



Fonte: Elaborada pelo autor.

uma cobertura de 99% dos extremos razoáveis dos dados.

**Interpretando os dados *BoxPlot*:** Esse gráfico fornece uma análise visual do conjunto de dados com as seguintes características: localização, dispersão, simetria, comprimento da cauda e valores aberrantes (*outliers*). O grande objetivo é verificar a distribuição dos dados. Assim, ao analisar um *BoxPlot* identifica-se os seguintes itens:

- Centro da distribuição dos dados: indicado pela linha mediana, ou Q2, no centro do quadrado.
- A dispersão dos dados: representada pelo intervalo interquartil (tamanho da caixa) que é a diferença entre o terceiro quartil (Q3) e o primeiro quartil (Q1), ou ainda pela amplitude que é a diferença entre o valor máximo (*max*) e o valor mínimo (*min*). Embora a amplitude seja de fácil entendimento, o intervalo interquartil é uma estatística mais robusta para medir variabilidade uma vez que não sofre influência de *outliers*.
- Simetria dos dados: O retângulo contém 50% dos valores do conjunto de dados. Uma distribuição simétrica terá a linha da mediana (Q2) no centro do retângulo. Quando a linha da mediana (Q2) está próxima ao primeiro quartil (Q1), os dados predominantemente possuem valores mais altos (positivamente assimétricos). E quando a posição da linha da mediana (Q2) está próxima ao terceiro quartil (Q3), os dados predominantemente possuem valores mais baixos (negativamente assimétricos).
- Os limites (*max*) e (*min*): As linhas que vão do retângulo até o maior valor considerado não discrepante.
- *Outliers*: indicam possíveis valores aberrantes. No *boxplot*, as observações são consideradas *outliers* quando estão abaixo ou acima do limite de detecção de *outliers*.

### 2.3.1.2 Regras de Associação

São técnicas de mineração de dados que representam combinações de itens que ocorrem com determinada frequência em uma base de dados. Essas regras foram apresentadas por (AGRAWAL; IMIELNISKI; SWAMI, 1993) em um estudo para descobrir as coocorrências entre conjuntos de itens a partir de um conjunto de dados estruturado como transações, nas quais cada transação contém um subconjunto de itens. Este estudo procurava encontrar relacionamentos entre os itens nas compras dos clientes numa visita ao supermercado e ficou conhecido como *market basket analysis*. Contudo, a aplicabilidade das regras de associação perpassa por diversas áreas distintas como análise epidemiológica, genética, recomendação de filmes, previsão de doenças cardíacas, ontologia, *bigdata*, computação ubíqua, Internet das Coisas (ROMÃO et al., 1999).

As regras de associação foram definidas em (AGRAWAL; IMIELNISKI; SWAMI, 1993) da seguinte forma: Sejam  $I = i_1, i_2, \dots, i_m$  um conjunto de  $m$  itens distintos e  $D$  uma base de dados formada por um conjunto de transações, onde cada transação  $T$  é composta por um conjunto de itens (*itemset*), tal que  $T \subseteq I$ . Uma regra de associação é uma expressão na forma  $A \Rightarrow B$ , onde  $A \subset I, B \subset I, A \neq \emptyset, B \neq \emptyset$  e  $A \cap B = \emptyset$ .  $A$  é denominado antecedente e  $B$  denominado consequente da regra. Tanto o antecedente, quanto o consequente de uma regra de associação pode ser formado por conjuntos contendo um ou mais itens. A quantidade de itens pertencentes a um conjunto de itens é chamada de comprimento do conjunto. Um conjunto de itens de comprimento  $k$  costuma ser referenciado como um  $k$ -*itemset* (GONSALVES, 2004).

A mineração de regras de associações utiliza medidas de interesse para gerar regras que satisfaçam as restrições iniciais do usuário. Existem várias medidas de interesse propostas pelos pesquisadores. Cada medida é utilizada, basicamente, para extrair um padrão específico dos dados (KUMAR; JOSHI et al., 2016). As medidas de interesse utilizadas nesta pesquisa são:

- **Suporte:** O suporte de um conjunto de itens  $Z$ ,  $Sup(Z)$ , representa a porcentagem de transações da base de dados que contêm os itens de  $Z$ . O suporte de uma regra de associação  $A \Rightarrow B$ ,  $Sup(A \Rightarrow B)$ , é dado por  $Sup(A \cup B)$ . Ou seja:

$$Sup(A \Rightarrow B) = (\text{Transações que contem } A \text{ e } B) / \text{Total de Transações}.$$

Se o limite mínimo de suporte escolhido for baixo, um grande número de regras será gerado e a avaliação de tais regras será complexa e demorada. No entanto se o valor mínimo de limite escolhido for alto, pode fazer com que o padrão seja ignorado e comprometer a tomada de decisão efetiva.

- **Confiança:** Já a confiança desta regra,  $Conf(A \Rightarrow B)$ , representa, dentre as transações que contêm  $A$ , a porcentagem de transações que também contêm  $B$ , é dada por  $Conf(A \Rightarrow B) = Sup(A \cup B) \div Sup(A)$ . Ou seja:



$\text{Conf}(A \Rightarrow B)$  = Número total de transações que contem A e B / número total de transações que contem A.

- **Lift:** uma das mais utilizadas para avaliar dependências. Dada uma regra de associação  $A \Rightarrow B$ , esta medida indica o quanto mais frequente torna-se B quando A ocorre. É dada por:

$$\text{Lift}(A \Rightarrow B) = \frac{\text{Conf}(A \Rightarrow B)}{\text{Sup}(B)}$$

- Se  $\text{Lift}(A \Rightarrow B) = 1$ , então A e B são independentes.
- Se  $\text{Lift}(A \Rightarrow B) > 1$ , então A e B são positivamente dependentes.
- Se  $\text{Lift}(A \Rightarrow B) < 1$ , então A e B são negativamente dependentes.

Esta medida varia entre 0 e  $\infty$ . e possui interpretação bastante simples: quanto maior o valor do *lift*, mais interessante a regra, pois A aumentou B numa maior taxa (GONSALVES, 2004).

- **Count:** Frequência da ocorrência de um conjunto de itens.
- **Coefficiente de correlação de Pearson ( $r$ ):** Mede o grau da correlação linear entre duas variáveis quantitativas. É um índice adimensional com valores situados entre -1,0 e 1,0 inclusive, que reflete a intensidade de uma relação linear entre dois conjuntos de dados.
  - Se  $r = 1$  Significa uma correlação perfeita positiva entre as duas variáveis.
  - Se  $r = -1$  Significa uma correlação negativa perfeita entre as duas variáveis. Ou seja, enquanto uma tem seu valor elevado a outra tem o seu valor reduzido.
  - Se  $r = 0$  Significa que as duas variáveis não dependem linearmente uma da outra. No entanto, pode existir uma outra dependência que seja "não linear". Assim, aconselha-se que o resultado  $r = 0$  seja investigado por outros meios.
- **Qui-quadrado ( $\chi^2$ ):** Teste para avaliar a correlação entre vários itens. Nas regras de associação é utilizado para testar a independência entre os itens das regras. Essencialmente, o teste  $\chi^2$  baseia-se na comparação das frequências observadas com as correspondentes frequências esperadas. Quanto mais próximas as frequências observadas estiverem das frequências esperadas, maior o peso da evidência em favor da independência. (KUMAR; JOSHI et al., 2016). O valor crítico da distribuição qui-quadrado com 1 grau de liberdade (tabela de contingência 2x2) em  $\alpha = 0,05$  é 3,84; quanto maior o valor do qui-quadrado, mais provável é a correlação das variáveis (HAN; PEI; KAMBER, 2011). O nível de significância é indicado pelo valor do *p-value*, fornecendo uma medida da força dos resultados de um teste, em contraste a uma simples rejeição ou não rejeição (WU et al., 2016).

Outras medidas de interesse poder ser utilizadas. Os trabalhos (TAN; KUMAR; SRIVASTAVA, 2002), (LENCA et al., 2007) e (GONSALVES, 2004) descrevem e detalham algumas das medidas de interesses.

As medidas de interesse objetivas são índices estatísticos utilizados para selecionar regras interessantes dentre as muitas que podem ser descobertas por um algoritmo de mineração de regras de associação. O suporte, confiança e *Lift* são exemplos de medidas de interesse objetivas. Já as medidas de interesse subjetivas levam em consideração o conhecimento, as intuições ou expectativas do usuário no momento da avaliação das regras. Em muitos casos uma regra de associação minerada é interessante para determinado usuário, mas não para outro. Fazendo com que a utilidade e a inesperabilidade sejam os fatores que podem tornar uma regra interessante. A medida de utilidade considera que uma regra é interessante se o usuário pode tirar proveito do padrão minerado. Já a medida de inesperabilidade considera que uma regra tem grande chance de ser interessante quando contradiz as expectativas do usuário, o que depende de suas convicções (GONSALVES, 2004).

## 2.4 Data Analytics

É a ciência de examinar dados brutos com o objetivo de encontrar padrões e tirar conclusões sobre essa informação, aplicando um processo algorítmico ou mecânico para obter informações. As técnicas de análise de dados podem revelar tendências e métricas que, de outra forma, seriam perdidas na massa de informações. Essas informações podem ser usadas para otimizar processos ou para aumentar a eficiência geral de um negócio ou sistema.

## 2.5 Transparência Traduzida

O site transparência traduzida, ver figura 7,<sup>1</sup> nasceu da ideia de publicizar, inicialmente, os resultados deste projeto de pesquisa. Dando ao cidadão uma visão simplificada dos dados públicos e brutos do sistema que controla os convênios de repasse federal (SICONV), obtidas pela análise e tradução da informação.

As informações publicadas no site são provenientes das análises e traduções dos dados públicos e brutos do sistema que controla os convênios de repasse federal (SICONV). Estas são traduzidas e publicadas separadamente por estado nos formatos de tabelas sumarizadas e gráficos. O acesso ao site não requer usuário nem senha, permitindo a qualquer cidadão navegar livremente pelas páginas do site.

Os dados existentes são referentes às transferências de recursos financeiros repassados pela União a estados, Distrito Federal, municípios e entidades privadas sem fins lucrativos, em

---

<sup>1</sup> [www.transparenciatraduzida.com.br](http://www.transparenciatraduzida.com.br)

Figura 7 – Logotipo do site Transparência Traduzida.



Fonte: Elaborada pelo autor.

decorrência da celebração de convênios, contratos de repasse ainda em execução e constituídos por emendas parlamentares propostas por parlamentares de forma individual.

# 3

## **Análise Inteligente de Dados Aplicada à Governança Pública e à Transparência de Convênios e Repasses Federais: Um Mapeamento Sistemático**

As seções a seguir descrevem o Mapeamento Sistemático utilizado para identificar e sistematizar as abordagens e técnicas computacionais de análise inteligente de dados que têm sido utilizadas para dar transparência, especificamente, aos repasses federais, ou para favorecer a Governança Pública em geral. O texto foi extraído do artigo escrito por Othon Stuart Ferreira Campos e Methanias Colaço Junior, respectivamente autor e orientador desta dissertação. O artigo foi publicado no 15<sup>th</sup> CONTECSI USP - *International Conference on Information Systems and Technology Management* - ISSN 1041-2448. Realizado na USP no período de 23 a 25 de maio, 2018 São Paulo, Brasil.

### **3.1 Método**

O processo de mapeamento sistemático da literatura (MSL) tem como objetivo fazer uma pesquisa em largura na literatura e não em profundidade ([KITCHENHAM, 2004](#)) e ([PETERSEN et al., 2008](#)). Por basear-se em um protocolo de pesquisa, o mapeamento pode ser reproduzido por outros pesquisadores. Os itens que constituem este protocolo são: objetivos de busca, palavras chave, filtros, mecanismos de busca acadêmica (MBA), critérios de exclusão/inclusão e dados a serem levantados. A definição dos objetivos de busca formaliza o que se pretende encontrar na literatura por meio do mapeamento. Os objetivos serão consequentemente refletidos nas questões, critérios, filtros e dados que serão levantados pela pesquisa. Neste sentido, o objetivo geral deste trabalho é "identificar as abordagens e técnicas computacionais de análise inteligente de dados que têm sido utilizadas para dar transparência, especificamente, aos repasses federais, ou para favorecer a Governança Pública em geral". Como repasses podem implicar presença de

valores que destoam da normalidade, bem como também podem representar desperdício público e associações suspeitas entre legisladores e entidades de todo tipo, na busca por pesquisas, foi dado um destaque à detecção de *outliers* e à descoberta de regras de associação.

### 3.1.1 Questões de pesquisa

Para alcançar o objetivo proposto por este mapeamento, foram elaboradas as seguintes questões de pesquisa:

- **Q1:** Quais são as abordagens de análise inteligente de dados mais utilizadas como base para controle e transparência de convênios e repasses federais, ou para favorecer a Governança Pública?
- **Q2:** Dentre as abordagens existentes, quais são as técnicas específicas utilizadas?
- **Q3:** Quais os métodos ou algoritmos específicos utilizados para descoberta de valores aberrantes em convênios e repasses federais, ou em dados gerais que favorecem a Governança Pública?
- **Q4:** No mesmo contexto, quais os métodos ou algoritmos específicos utilizados para análise de associações?
- **Q5:** Quais são os países com maior número de pesquisadores publicando sobre esse tema?
- **Q6:** Quais os anos que tiveram mais publicações nessa área?
- **Q7:** Quais os principais periódicos e conferências sobre o tema?
- **Q8:** Quais os meios de publicação mais populares?

### 3.1.2 Escopo da pesquisa

Para operacionalizar o MSL, foi acessada a base *Scopus* por meio do portal de periódicos da CAPES (<https://www.periodicos.capes.gov.br>), pois ele permite fazer download dos artigos sem restrições. A *Scopus* foi escolhida por incluir buscas em diferentes bancos de dados científicos (IEEE, ACM, Springer e Elsevier). As fontes foram selecionadas de acordo com a disponibilidade de consulta através da internet, as quais foram indexadas nas bases citadas, podendo ser encontradas por meio da busca por palavras-chave. A respeito do idioma, foram selecionados apenas trabalhos em inglês. A respeito da área, foram selecionados apenas trabalhos referentes à Ciência da Computação. A respeito do tipo de publicação, foram selecionados apenas artigos publicados em conferências ou periódicos.

### 3.1.3 Método de Busca de Publicações

As fontes foram acessadas pela internet. A criação das palavras-chave se baseou na definição do objeto (computação inteligente), da ação (análise de dados) e da área de aplicação (Governança Pública e Repasses Federais). Com a utilização destas definições em inglês e do uso de vários sinônimos, associados ao pressuposto de que os estudos primários estariam contidos nas áreas da computação que lidam com inteligência e análise de dados, a *String* de busca utilizada na base *Scopus* foi:

```
( "transfer* public resourc*" OR "public covenant*" OR "Public transpa-  
rency" OR "Public Governance" ) AND ( "data mining" OR "data analy-  
tics" OR "outlier" OR "association rules" OR "big data" OR "business  
intelligence" OR "data science" OR "artificial intelligence" OR "machine lear-  
ning" ) AND PUBYEAR > 2007 AND ( LIMIT-TO ( SUBJAREA , "COMP  
" ) ) AND ( LIMIT-TO ( DOCTYPE , "cp " ) OR LIMIT-TO ( DOCTYPE , "ar  
" ) ) AND ( LIMIT-TO ( LANGUAGE , "English " ) )
```

### 3.1.4 Critérios de Seleção

Para filtrar os artigos relevantes para este Mapeamento Sistemático, foram estabelecidos os critérios de inclusão e exclusão dos mesmos. O estudo incluiu os seguintes critérios de inclusão:

- O resultado deve conter o tema deste estudo no título, resumo ou palavras-chave;
- O resultado deve estar disponível para consulta online;
- O resultado precisa explorar um algoritmo, técnica, mecanismo ou abordagem de análise inteligente de dados utilizada para controle e transparência de convênios e repasses federais, ou para favorecer a Governança Pública.

A confirmação dos critérios de inclusão foi dada pela análise do resumo e da introdução de cada artigo. Em paralelo, os artigos foram analisados com relação aos critérios de exclusão. Os critérios de exclusão descritos abaixo também foram aplicados:

- Artigos que não dizem respeito ao campo da ciência da computação;
- Estudos secundários, pois eles tratam de abordagens de terceiros;
- Artigos que estavam indisponíveis;
- Artigos com ano de publicação anterior a 2008;
- Estudos preliminares.

Após a aplicação dos critérios de seleção, os trabalhos foram avaliados. Dentre os 122 encontrados, 6 foram selecionados para compor os estudos primários. Após a seleção, os estudos primários foram referenciados para leitura e análise.

## 3.2 Ameaças à Validade

As ameaças à validade deste estudo foram:

- **Validade de Construção:** A *String* de busca e as questões de pesquisa utilizadas podem não cobrir a área de análise inteligente de dados para Governança Pública e para controle de repasses federais. Para mitigar esta ameaça, procuramos construir a *string* mais abrangente possível, quanto aos termos que pudessem ser usados na área, além de usar três pesquisadores e um artigo de controle.
- **Validade Interna (Extração de dados):** Como três pesquisadores foram responsáveis por classificar e extrair os principais algoritmos de cada publicação, vieses ou problemas na extração dos dados podem ameaçar a validade da caracterização dos dados; (Viés de Seleção): No começo do estudo, os artigos foram incluídos ou excluídos no mapeamento de acordo com julgamento dos próprios pesquisadores. Isso significa que alguns estudos podem ter sido categorizados incorretamente. Para mitigar estas ameaças, as revisões da seleção e extração foram feitas por todos os pesquisadores envolvidos, com uma votação final para as discordâncias.
- **Validade Externa:** Apesar da *Scopus* ser a maior base de literatura científica, com mais de 21,500 *journals* e 60 milhões de registros (SCOPUS, 2017), não é possível afirmar que os resultados desse mapeamento abrangeram toda a área da ciência da computação. No entanto, este trabalho apresentou evidências das principais técnicas utilizadas, identificando lacunas a serem exploradas e servindo como guia para futuros trabalhos nesta linha.

## 3.3 Discussão

Nesta seção, são apresentados os resultados da análise dos estudos primários, respondendo às questões de pesquisa apresentadas anteriormente. Após a execução do protocolo de pesquisa para o mapeamento sistemático de literatura, foram encontrados 122 trabalhos, os quais, após a aplicação dos critérios de seleção, foram reduzidos para 6 artigos. A Tabela 3 mostra, de forma sintética, as fontes com seus respectivos números de artigos retornados pela *String* de busca, bem como a quantidade de artigos selecionados por fonte.

Na Tabela 4, estão listados os 6 estudos primários selecionados por este estudo, bem como seus respectivos números de referência.

Tabela 3 – Processo de seleção.

Fontes	Estudos retornados pela String	Estudo selecionado
Springer	42 (34%)	3 (50%)
IEEE Computer Society	11 ( 9%)	2 (33%)
ACM	19 (16%)	1 (17%)
Elsevier	6 (5%)	( 5%)
Inderscience Enterprises Ltd.	6 ( 5%)	0
IOS Press	6 ( 5%)	0
Taylor & Francis.	5 ( 4%)	0
ACPI	4 ( 3%)	0
Emerald Group Publishing Ltd.	4 ( 3%)	0
Outros	19 (16%)	0

Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas à partir da consulta na base *Scopus*.

Tabela 4 – Estudos primários selecionados.

Título	Referência
Business intelligence systems as management, accountability and transparency tools for the government: The case of platform aquarius	(Capuano, 2014)
Are the methodologies for producing linked open data feasible for public administrations?	(Boselli, 2014)
Transparency dimensions of big and open linked data: Transparency as being synonymous with accountability and openness	(Matheus R. &., 2015)
Innovating good regulatory practice using mixed-initiative social media analytics and visualization	(Lemieux, 2016)
Increasing open government data transparency with spatial dimension	(Dessì, 2016)
Can big data improve public policy analysis?	(Shi, 2017)

Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas à partir da consulta na base *Scopus*.

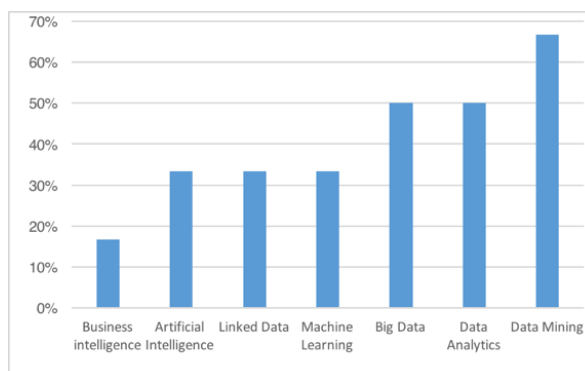
Para responder às questões, devemos levar em consideração que não foram localizados, em qualquer contexto, estudos específicos para transparência de convênios e repasses federais.

À questão de pesquisa **Q1** (Quais são as abordagens de análise inteligente de dados mais utilizadas como base para controle e transparência de convênios e repasses federais, ou para favorecer a Governança Pública?) destacamos *Data Mining* como a mais utilizada, perfazendo 67% dos estudos primários. Na sequência, *Data Analytics* e *Big Data* aparecem em 50% dos



estudos, bem como *Machine Learning*, *Artificial Intelligence* e *Linked Data* aparecem com 33% e *Business intelligence*, com 17%. A figura 8 mostra a distribuição do percentual diante dos temas abordados.

Figura 8 – Distribuição dos temas abordados.



Fonte: Elaborada pelo autor (2018).

A resposta para a questão **Q2** (Dentre as abordagens existentes, quais são as técnicas específicas utilizadas?) está demonstrada na Tabela 5 os dois únicos algoritmos citados, explicitamente, em um único trabalho. Todos os outros não detalham as técnicas utilizadas na análise dos dados.

Tabela 5 – Técnicas específicas.

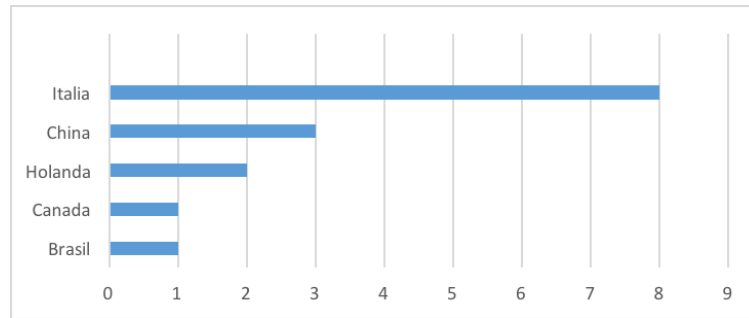
Técnicas	Artigo
Ridge Regression	(Shi, 2017)
Elastic Net	(Shi, 2017)

Fonte: Elaborada pelo autor (2018).

Já para as questões **Q3** (Quais os métodos ou algoritmos específicos utilizados para descoberta de valores aberrantes em convênios e repasses federais, ou em dados gerais que favorecem a Governança Pública?) e **Q4**: (No mesmo contexto, quais os métodos ou algoritmos específicos utilizados para análise de associações?), não foi identificada a utilização de algoritmos para descoberta de valores aberrantes, tampouco para descoberta de regras de associação.

A questão **Q5** (Quais são os países com maior número de pesquisadores publicando sobre esse tema?) o mapeamento contabilizou 15 autores nos 6 estudos primários. Os pesquisadores têm origem em 5 países diferentes, como mostrado na Figura 9. A liderança da Itália pode representar a maturidade adquirida desde a década de 90, decorrente da operação “Mãos Limpas”, responsável pela descoberta de licitações irregulares e o uso do poder público em benefício particular e de partidos políticos. A mesma operação comprovou ainda que empresários pagavam propinas a políticos para vencer licitações de construção de ferrovias, autoestradas, prédios públicos, estádios e obras civis em geral.

Figura 9 – Periódicos e Conferências.



Fonte: Elaborada pelo autor.

Para a questão **Q6** (Quais os anos que tiveram mais publicações nessa área?) observamos que os anos com maior número de publicações foram 2014 e 2016, com 2 artigos cada. Os anos de 2015 e 2017 aparecem com apenas uma publicação. Nos anos que antecedem o ano de 2014, não foram publicados artigos. É notória a escassez de publicações sobre a área.

A resposta para a **Q7** (Quais os principais periódicos e conferências sobre o tema?) se encontra na Tabela 6.

Tabela 6 – Periódicos e Conferências.

Veículos	Total de Publicações	Publisher	Journal	Proceeding
ACM International Conference Proceeding Series	1	ACM		X
Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)	2	Springer		X
International Conference on Data Management Technologies and Applications	1	Springer		X
International Conference for E-Democracy and Open Government, CeDEM	1	IEEE		X
International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises, WE-TICE	1	IEEE		X

Fonte: Elaborada pelo autor (2018).

Por fim, a resposta da questão **Q8** (Quais os meios de publicação mais populares?) observamos que 100% das publicações para trabalhos sobre o tema são as conferências.

## 3.4 Considerações Finais do Capítulo

Como não foi encontrado nenhum outro trabalho similar de Mapeamento ou Revisão Sistemática, supõe-se que se trate do primeiro trabalho deste tipo nesta área específica do conhecimento científico. Como resultados, para a área de Governança Pública, dentre os raros trabalhos encontrados, identificou-se que as abordagens mais citadas foram: *Data Mining*, em 67% dos trabalhos, seguida por *Data Analytics* e *Big Data*, ambas com 50%, *Machine Learning*, *Artificial Intelligence* e *Linked Data*, com 33%, e, por fim, *Business intelligence*, em apenas um trabalho (17%). Apenas dois artigos citaram *Machine Learning* e *Artificial Intelligence*, o que denota que suas aplicações ainda são pouco exploradas. Isto sugere que estas áreas podem representar um bom campo para trabalhos futuros sobre o assunto. Somente um trabalho elucidou os algoritmos (*Ridge Regression* e *Elastic Net*) utilizados para fazer a análise dos dados, em outras palavras, é o único artigo que realmente evidencia o uso prático de computação inteligente para transparência pública.

Apesar de supor-se o contrário, não foi possível identificar métodos ou algoritmos para a descoberta de desvios, tampouco para identificação de regras de associação relacionadas aos convênios e repasses federais, ou à Governança Pública. Uma outra evidência de tema pouco explorado, representando uma oportunidade na áreas de mineração de dados para trabalhos futuros. Acredita-se que este trabalho é relevante para a academia, para os governos e para a comunidade em geral, apresentando-lhes tendências e ausências no uso de inteligência para transparência pública. Além disso, ele pode ofertar mais uma capilaridade na busca pelas melhores soluções para um cenário da política adotada pelo Brasil.

# 4

## Estudo de caso de descrição dos Dados de Convênios Brasileiros

Esta seção apresenta o estudo de caso realizado com a mineração de dados brutos na base de dados aberta do SICONV, disponível para *download* no ([SICONV, 2016](#)). Por conveniência, foi escolhido estudar o valor global dos convênios (uma vez que este valor representa o total de gastos a ser disponibilizado para o convênio), ainda em execução, firmados entre o Governo Federal e órgãos/entidades ou entes federados ou ainda com entidades do setor privado. O valor global do convênio corresponde ao valor do repasse mais o valor da contrapartida.

### 4.1 Definição do objeto

O objetivo deste estudo de caso é avaliar valores aberrantes (*outliers*) nos valores globais dos convênios, ainda em execução, a fim de identificar convênios que apresentem uma grande variação ou inconsistência em relação aos demais valores globais. Para este estudo, foram disponibilizados, *on-line*, no site [www.transparenciatraduzida.com.br](http://www.transparenciatraduzida.com.br), dados de todo o território nacional. Nesta dissertação, detalhamos a análise dos dados do Estado de Sergipe, no período de 2008 a 2018.

### 4.2 Planejamento

A questão de pesquisa que precisa ser respondida por meio do estudo de caso é: Existem discrepâncias entre os valores dos convênios firmados no estado de Sergipe? A estudo foi realizado em maio de 2018, com os dados brutos mais recentes, no momento da pesquisa, disponibilizados para *download*. Como os dados são abertos e estão disponíveis a todos os cidadãos, não existiu a necessidade de reuniões com as entidades envolvidas. Os estudo foi realizado diretamente pelos pesquisadores.

### 4.2.1 Seleção de contexto

Foram selecionados todos os convênios ainda em execução, no período de 2008 a 2018, propostos por emendas parlamentares individuais. Os dados analisados foram obtidos no portal de convênios ([SICONV, 2016](#)), sendo selecionados os dados dos arquivos referentes aos convênios, emendas parlamentares e propostas cadastradas. Para atender à questão de pesquisa, foram analisados os convênios do estado de Sergipe.

### 4.2.2 Instrumentação

O processo de instrumentação teve início com a configuração do ambiente para o estudo de caso e planejamento de coleta de dados. Os materiais / recursos utilizados foram:

- Arquivos com os dados do SICONV. Mais especificamente: `siconv_convenio.csv`, `siconv_conv_programa_proposta.csv`, `siconv_proposta.csv` e `siconv_emenda.csv`;
- Banco de Dados PostgreSQL, versão 9.2;
- Arquivos com o *script* de criação do projeto de banco de dados;
- Software *Talend Open Studio*, responsável pela construção e execução do ETL de povoamento do banco de dados;
- Ambiente de execução do Java. Versão igual ou superior a 1.7;
- Software livre R, para análises de *outliers*, criação dos gráficos comuns e boxplots.

## 4.3 Preparação

A preparação do ambiente consistiu na instalação do ambiente de execução Java, instalação do PostgreSQL, instalação do ambiente R e do *Talend Open Studio*. Além da instalação dos *softwares* citados, foi necessário efetuar o *download* dos arquivos no portal de convênios ([SICONV, 2016](#)). O *software Talend Open Studio* foi utilizado para construir o ETL (Extração, Transformação e Carga) de povoamento da base de dados. Também foi desenvolvido um projeto de banco de dados (ver apêndice [A](#)), utilizando a linguagem de manipulação e criação de dados SQL, o qual também foi utilizado pelo ETL para geração de resultados exportados para arquivos CSV (*Comma Separated Values*), contendo os dados necessários para análise de *outliers*.

## 4.4 Execução

Depois de feito o *download* dos arquivos, foi criado o banco de dados para receber os dados dos arquivos. Depois do banco criado, foi executado o ETL, no ambiente do *software Talend Open Studio*, povoando o banco de dados. Em seguida, foram executados os comandos, (ver apêndice B), criados no projeto de banco de dados, responsáveis pela criação do arquivo CSV, contendo os valores a serem estudados. Por fim, o arquivo CSV foi exportado para o ambiente do R, onde foram construídos os gráficos do tipo *boxplots* e feita as análises dos valores aberrantes (*outliers*).

### 4.4.1 Coleta de dados

Para coleta de dados e evidências, o ETL selecionou os dados desejados e os armazenou em um arquivo do tipo CSV. Este arquivo, por sua vez, fora exportado para o ambiente do R, de onde foram retiradas as informações para gerar os gráficos e valores aberrantes (*outliers*).

### 4.4.2 Detecção de Outliers

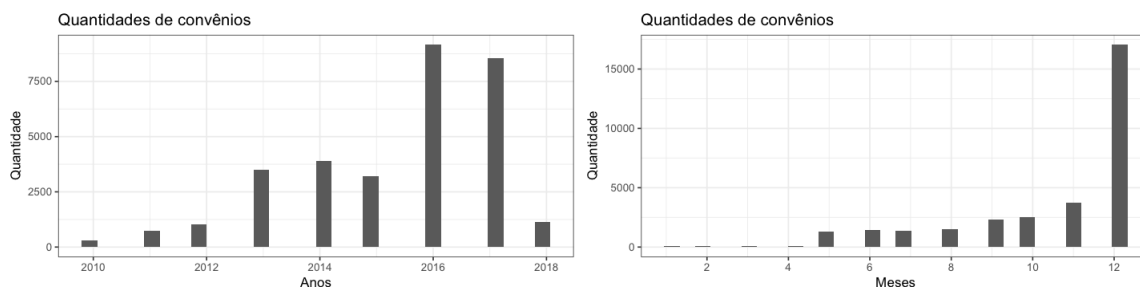
O conjunto de dados analisado possui um total de 32.855 convênios, ainda em execução, em todo o país, com 72 duas variáveis em cada um. Para o Estado de Sergipe, esse valor é de 850 convênios. Para a identificação de possíveis valores aberrantes, foram escolhidas as variáveis: ano, mês, UF do proponente do convênio e o valor global do convênio. Foram gerados gráficos para identificação de valores aberrantes para todo o Brasil e para o estado de Sergipe. No entanto, para este estudo, as análises foram feitas apenas para o estado de Sergipe. Dados de outros estados da nação estão publicados no site [www.transparenciatraduzida.com.br](http://www.transparenciatraduzida.com.br).

Para a detecção dos valores aberrantes (*outliers*), utiliza-se o valor para o coeficiente no cálculo dos limites **mim.** e **max.**, dando uma cobertura de 99% dos dados analisados.

## 4.5 Análise e interpretação

Para iniciar a análise, começaremos pela distribuição das quantidades de convênios, mostrada na figura 10, em todo o Brasil, por ano e por mês. Nota-se um leve incremento na quantidade de convênios em execução até o ano de 2012. Nos anos de 2013 e 2016, nota-se um grande salto, incremento de aproximadamente 300%, no número de convênios em execução, em relação ao ano anterior. Apenas o ano de 2015 registrou um decremento no número de convênios em execução, quando comparado ao ano anterior. O ano de 2016 claramente se destaca, com aproximadamente 10.000 convênios. O ano de 2017, com aproximadamente 8.500 convênios, também merece destaque. O ano de 2018, apesar de possuir informações até o mês de maio, momento que fora executado o estudo de caso, já supera a quantidade de convênios do ano de 2012.

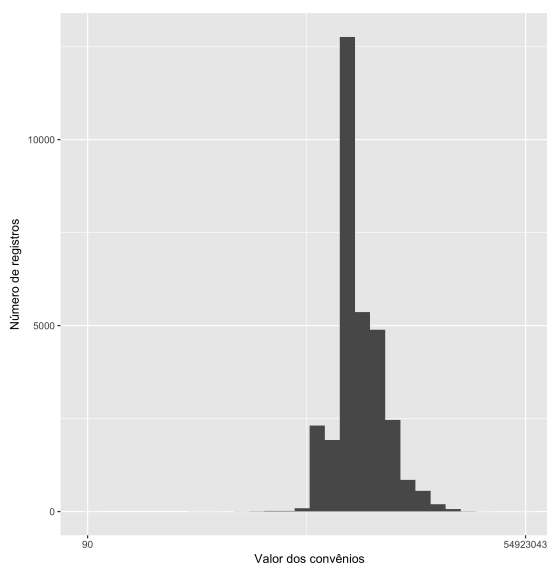
Figura 10 – Distribuição dos convênios de todo o Brasil.



Fonte: Elaborada pelo autor (2018).

Já no gráfico que mostra a distribuição de convênios por mês, observa-se que o mês de dezembro concentra um valor expressivo na quantidade de convênios, sendo superior ao somatório das quantidades dos meses anteriores. Em relação aos valores dos convênios, observamos, na figura 11, que é predominante o número de convênios em execução com valores mais elevados (assimetria positiva).

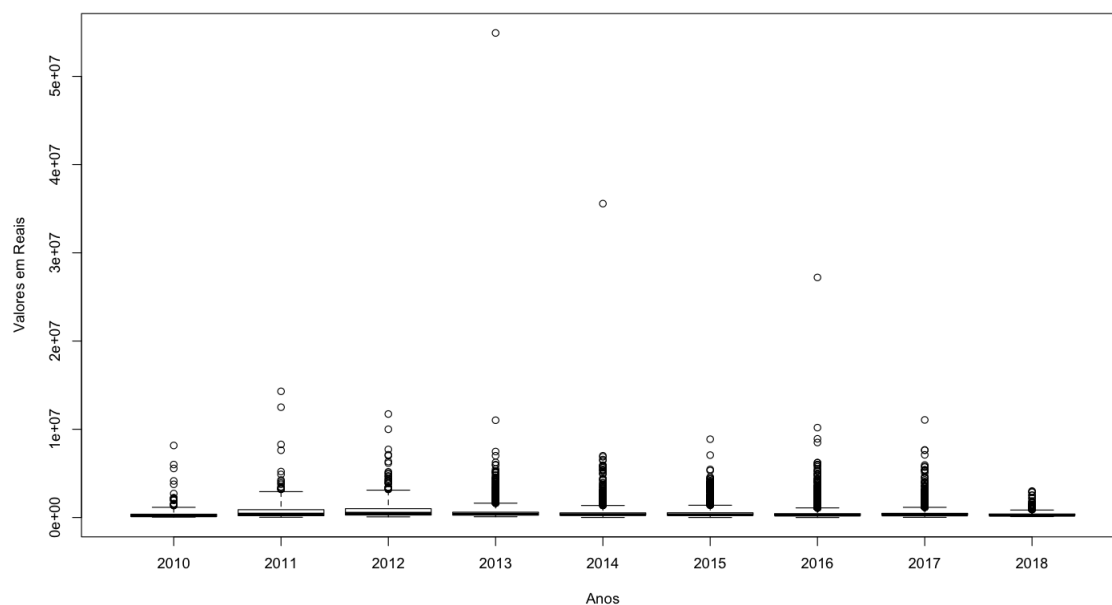
Figura 11 – Distribuição dos valores dos convênios de todo o Brasil.



Fonte: Elaborada pelo autor (2018).

Analisando o gráfico mostrado na figura 12, ver 2.3.1.1.1, pode-se observar alguns valores aberrantes nos dados dos convênios de todo o país, ao ponto de não conseguirmos detalhar os valores referentes aos quartis, ou mesmo os valores mínimos e máximos. No entanto, como esse gráfico está exibindo os dados dos convênios de todo o país, os valores exibidos como aberrantes podem ser considerados apenas possíveis outliers, uma vez que devem ser analisados dentro da realidade da unidade federativa com a qual o convênio foi firmado.

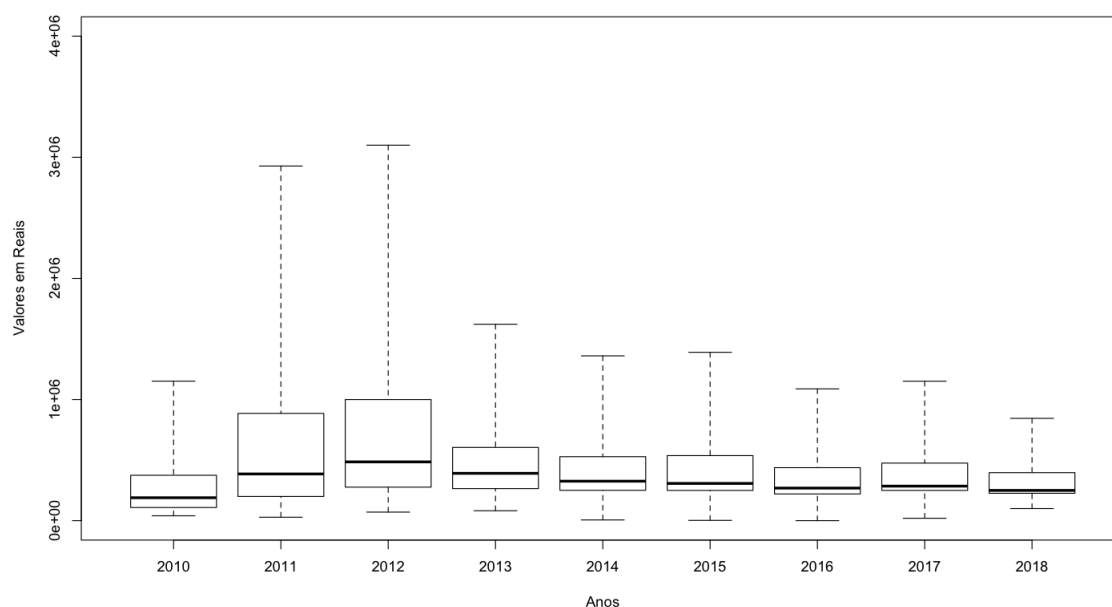
Para um melhor entendimento da dispersão dos valores dos convênios nos diferentes anos, os valores aberrantes (*outlier*) foram ocultados na figura 13. Neste gráfico, pode-se observar

Figura 12 – *Outliers* dos convênios de todo o Brasil.

Fonte: Elaborada pelo autor (2018).

uma predominância de uma maior quantidade de convênios com valores menores, caracterizando uma assimetria positiva dos dados (ver 2.3.1.1.1).

Figura 13 – Dispersão dos valores dos convênios ativos no Brasil.



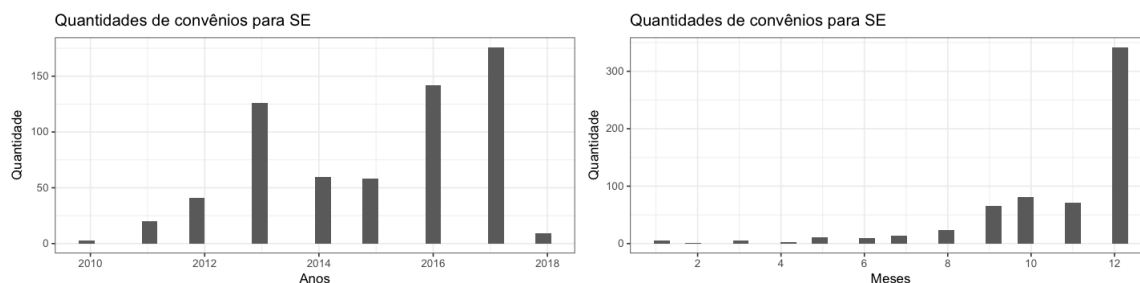
Fonte: Elaborada pelo autor (2018).

Restringindo os dados para o Estado de Sergipe, no intuito de atender à questão de pesquisa, a figura 14 mostra que, assim como nos dados nacionais, os anos com maiores incrementos nos números de convênios, em relação ao ano anterior, são 2013 e 2016. No entanto, o ano de 2017 manteve a tendência de alta, contrariando os números dos convênios firmados



em todo o Brasil. Com relação à distribuição de convênios firmados por mês, Sergipe também possui o maior número no mês de dezembro, mantendo assim a mesma tendência dos números nacionais.

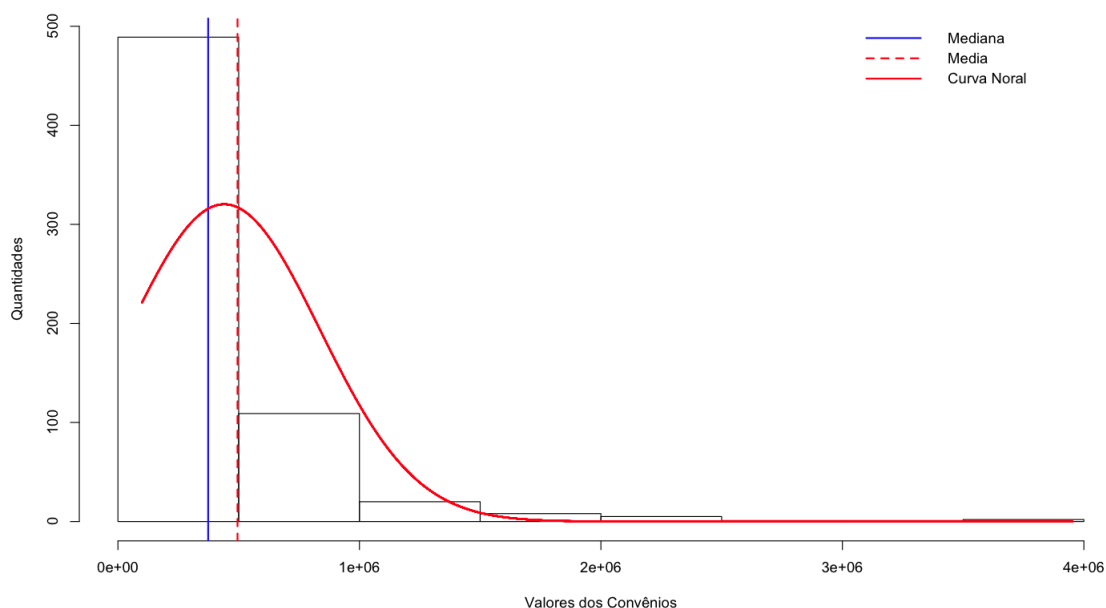
Figura 14 – Distribuição dos convênios no Estado de Sergipe.



Fonte: Elaborada pelo autor (2018).

Em relação à distribuição de valores, o histograma representado pela figura 15 nos indica que o maior número de convênios possuem os menores valores (assimetria positiva). A barra contendo o pico (moda) representa o conjunto de valores mais comuns.

Figura 15 – Distribuição de valores dos convênios no Estado de Sergipe.



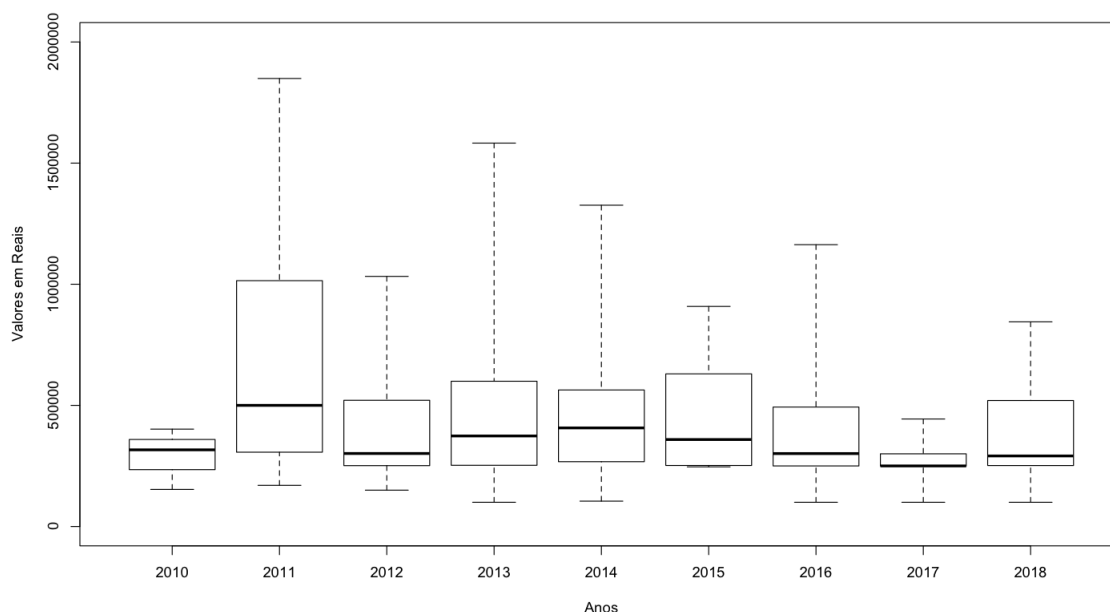
Fonte: Elaborada pelo autor (2018).

No gráfico representado na figura 16,<sup>1</sup> os valores aberrantes (*outliers*) foram ocultados a fim de entender melhor a dispersão dos valores dos convênios nos diferentes anos. Também, pode-se observar uma assimetria positiva nos valores dos convênios. Observou-se também que, na maioria dos anos, 50% dos valores ficam entre R\$ 250.000,00 e R\$ 500.000,00, aproximadamente. Excetuando-se o ano de 2011, que possui 50% com valores entre R\$ 400.000,00 e

<sup>1</sup> ver seção 2.3.1.1.1, para maiores detalhes sobre este tipo de gráfico

R\$ 1.000.000,00. Em relação ao valor da mediana (Q2), em todos os anos, exceto 2010 e 2014, está próxima ao primeiro interquartil (Q1), corroborando com a evidência de que os dados, predominantemente, possuem valores mais baixos. No ano de 2010, a mediana se aproxima do terceiro interquartil (Q3), denotando que a maior parte dos convênios possuem valores maiores. Já o ano de 2014, o valor da mediana está próximo ao centro no interquartil, mostrando um equilíbrio na distribuição normal dos valores dos convênios.

Figura 16 – Distribuição dos valores dos convênios no Estado de Sergipe.



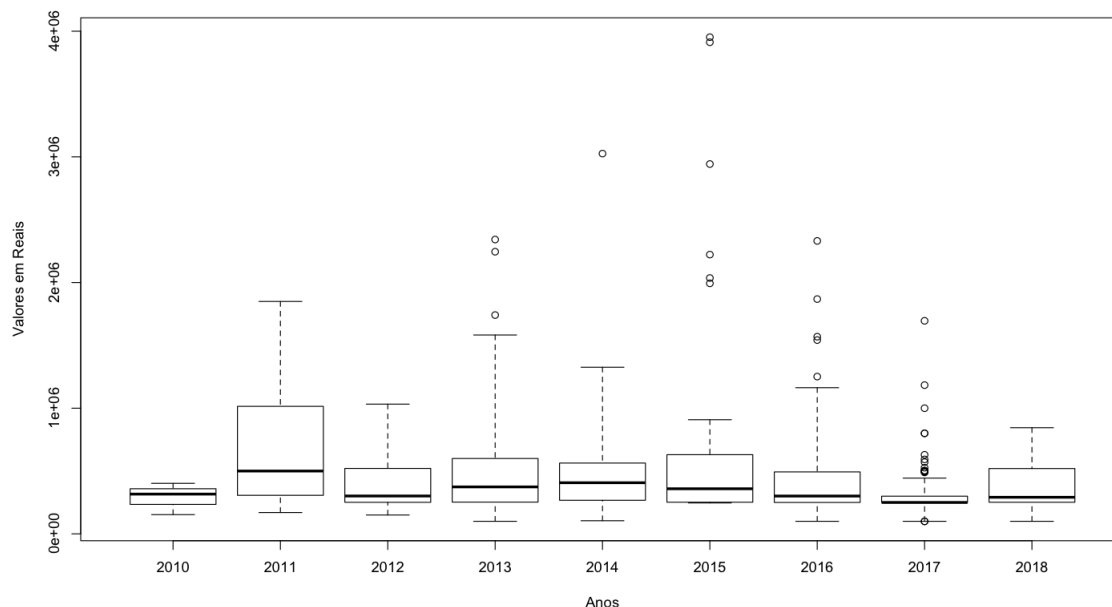
Fonte: Elaborada pelo autor (2018).

Analisando o gráfico, mostrados na Figura 17,<sup>2</sup> pode-se observar alguns valores discrepantes (*outlier*) nos dados dos convênios do Estado de Sergipe. Os anos de 2010, 2011, 2012 e 2018 não apresentaram valores aberrantes (*outlier*) e demonstram uma assimetria positiva, significando que o valor da maioria dos convênios são relativamente menores em relação ao número de convênios com valores mais elevados. Já os anos de 2013, 2014, 2015, 2016 e 2017 apresentam valores discrepantes. O ano de 2013 possui 3 *outliers*, 2014 apresentou apenas 1 *outlier*, 2015 possui 6 *outliers*, 2016 com 5 *outliers* e 2017, ano que apresentou o maior número, com 21 *outliers*, sendo apenas dois abaixo do limite mínimo (*mim*). O maior valor aberrante foi encontrado no ano de 2015, um convênio no valor de R\$ 3.952.638,90 (três milhões, novecentos e cinquenta e dois mil, seiscentos e trinta e oito reais e noventa centavos). Este convênio ainda está ativo e foi viabilizado por uma emenda parlamentar da Senadora Maria do Carmo, para a área da Saúde. O menor valor aberrante foi encontrado no ano de 2017, no valor de R\$ 99.562,71 (noventa e nove mil, quinhentos e sessenta e dois reais e setenta e um centavos). Esse convênio ainda está ativo e foi viabilizado por uma emenda parlamentar do Senador Eduardo Amorim, para a área da Cultura. Todos os convênios com valores aberrantes, no Estado de Sergipe, no

<sup>2</sup> ver seção 2.3.1.1.1 para maiores detalhes sobre este tipo de gráfico.

período de 2008 a 2018, estão listados no Anexo B. Essas informações também podem ser obtidas no site Transparência Traduzida, como demonstrado pela Figura 18<sup>3</sup>.

Figura 17 – *Outliers* dos convênios no Estado de Sergipe.



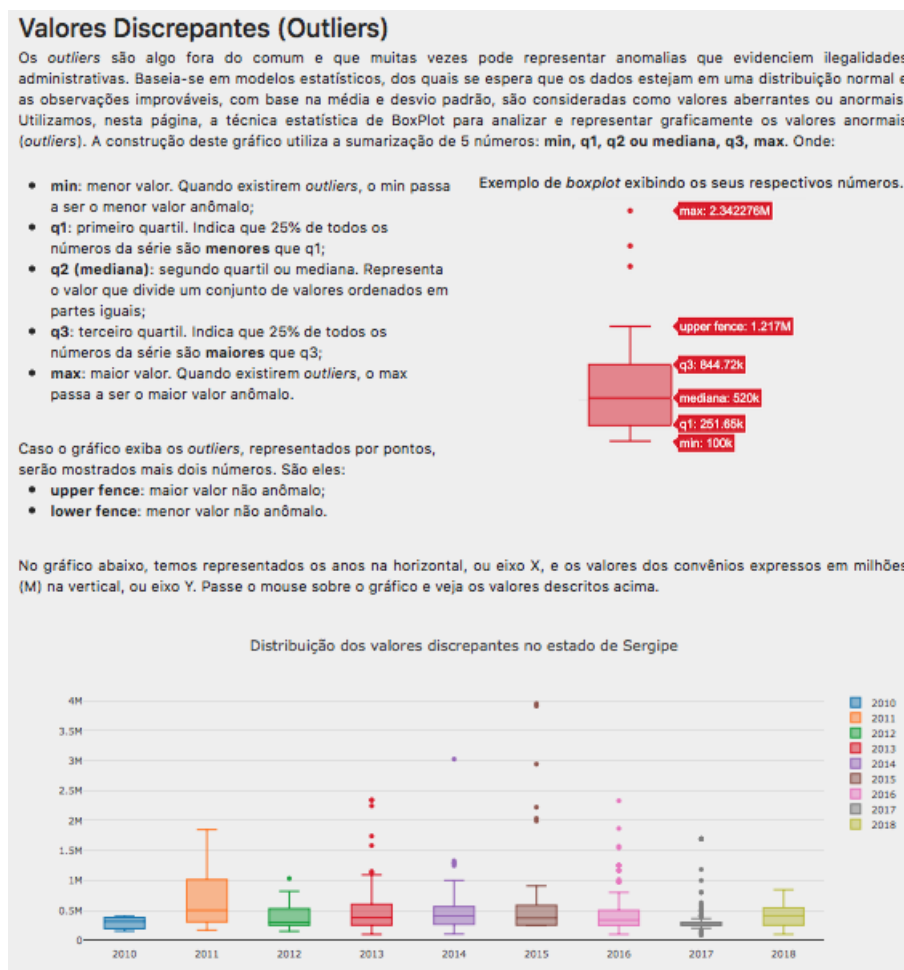
Fonte: Elaborada pelo autor (2018).

No total, foram encontrados 36 convênios com valores aberrantes *outliers*, atendendo a 20 diferentes proponentes, 8 áreas distintas, 16 parlamentares, 34 favorecidos, bem como totalizado um valor de R\$ 47.961.892,70 (Quarenta e sete milhões e novecentos e sessenta e um mil e oitocentos e noventa e dois reais e setenta centavos).

A tabela 7 exibe os 10 proponentes com os maiores valores de convênios com valores discrepantes. Observa-se que o proponente "Estado de Sergipe", concentra 11 convênios, 30,56% da quantidade total, com um valor de R\$ 16.273.842,35 (Dezesseis milhões e duzentos e setenta e três mil e oitocentos e quarenta e dois reais e trinta e cinco centavos). Logo em seguida, aparece o Município de Aracaju, com 3 convênios, 11,67% da quantidade total, com um valor de R\$ 5.598.769,74 (Cinco milhões e quinhentos e noventa e oito mil e setecentos e sessenta e nove reais e setenta e quatro centavos). Em terceiro lugar, aparece o Município de Capela, com 2 convênios, 8,51% da quantidade total, com um valor de R\$ 4.083.525,85 (Quatro milhões e oitenta e três mil e quinhentos e vinte e cinco reais e oitenta e cinco centavos). Apenas esses três primeiros proponentes concentram 54,12%, ou R\$ 25.956.137,94 (Vinte e cinco milhões e novecentos e cinquenta e seis mil e cento e trinta e sete reais e noventa e quatro centavos), do valor total dos convênios com valores discrepantes (*outliers*).

Na tabela 8, são exibidas as 8 áreas que atendem os convênios. Observa-se que apenas o Ministério das Cidades é responsável por 15 convênios com valores discrepantes, 41,67% do

<sup>3</sup> imagem retirada do site [www.transparenciatraduzida.com.br](http://www.transparenciatraduzida.com.br)

Figura 18 – *Outliers* no site Transparência Traduzida

Fonte: [www.transparenciatraduzida.com.br](http://www.transparenciatraduzida.com.br).

total de convênios, totalizando R\$ 18.372.604,96 (Dezoito milhões e trezentos e setenta e dois mil e seiscentos e quatro reais e noventa e seis centavos), o que representa 38,31% do valor total. O Ministério dos Esportes aparece em segundo lugar, com 7 convênios, 18,65% do total de convênios, totalizando R\$ 8.946.408,76 (Oito milhões e novecentos e quarenta e seis mil e quatrocentos e oito reais e setenta e seis centavos), o que representa 18,65% do valor total. Apenas esses dois Ministérios concentram 56,96%, ou seja, R\$ 27.319.013,72 (Vinte e sete milhões e trezentos e dezenove mil e treze reais e setenta e dois centavos) do valor total dos convênios, com valores aberrantes. O Ministério do Turismo e o Ministério do Desenvolvimento Social possuem os menores valores para os convênios com discrepância nos valores, R\$ 526.407,18 (Quinhentos e vinte e seis mil e quatrocentos e sete reais e dezoito centavos) e R\$ 500.000,00 (Quinhentos mil reais), respectivamente.

Em relação aos parlamentares, a tabela 9 exibe os 10 parlamentares com maiores valores de emendas. Os três primeiros são Almeida Lima, João Daniel e Maria do Carmo Alves. Todos eles com valores superiores a 6 milhões de reais. Apenas esses três parlamentares juntos concentram 44,71% do valor total dos convênios que possuem valores discrepantes *outlier*, o que totaliza

Tabela 7 – Quantidades e valores dos convênios por proponente.

<b>Proponente</b>	<b>Quant.</b>	<b>%</b>	<b>Valor Total</b>	<b>%</b>
Estado de Sergipe	11	30,56%	R\$ 16.273.842,35	33,93%
Mun. de Aracaju	3	8,33%	R\$ 5.598.769,74	11,67%
Mun. de Capela	2	5,56%	R\$ 4.083.525,85	8,51%
Fundo Mun. de Saúde	1	2,78%	R\$ 3.952.638,90	8,24%
Mun. de Frei Paulo	1	2,78%	R\$ 3.026.130,00	6,31%
Ass. Hospitalar de Sergipe	1	2,78%	R\$ 2.943.000,00	6,14%
Mun. de Itabaiana	1	2,78%	R\$ 2.331.252,34	4,86%
Mun. Rosário do Catete	1	2,78%	R\$ 2.245.235,86	4,68%
Mun. de Aquidabã	2	5,56%	R\$ 1.429.147,05	2,98%
Mun. de N. S. da Glória	2	5,56%	R\$ 1.000.000,00	2,08%

Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

Tabela 8 – Quantidades e valores dos convênios por área.

<b>Área</b>	<b>Quant.</b>	<b>%</b>	<b>Valor Total</b>	<b>%</b>
Min. das Cidades	15	41,67%	R\$ 18.372.604,96	38,31%
Min. do Esporte	7	19,44%	R\$ 8.946.408,76	18,65%
Min. da Saúde	3	8,33%	R\$ 7.395.638,90	15,42%
Min. da Integração Nacional	2	5,56%	R\$ 6.134.687,00	12,79%
Min. da Cultura	6	16,67%	R\$ 4.217.686,61	8,79%
SEAD	1	2,78%	R\$ 1.868.459,29	3,90%
Min. do Turismo	1	2,78%	R\$ 526.407,18	1,10%
Min. do Desen. Social	1	2,78%	R\$ 500.000,00	1,04%

Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

R\$ 20.205.216,87 (Vinte milhões e duzentos e cinco mil e duzentos e dezesseis reais e oitenta e sete centavos). O parlamentar Almeida Lima possui apenas duas emendas parlamentares, no entanto, estas totalizam o maior valor entre os parlamentares. O Senador Eduardo Amorim possui 7 emendas, sendo o parlamentar com maior número de emendas. Porém, o somatório dos valores das suas emendas totaliza R\$ 4.051.966,61 (Quatro milhões e cinquenta e um mil e novecentos e sessenta e seis reais e sessenta e um centavos).

No que se refere aos favorecidos, identificou-se um total de 34 entidades que receberam pagamentos pela execução dos projetos dos convênios. Desta quantidade, a tabela 10 lista os 10 favorecidos com maiores valores recebidos. Apesar do valor total dos convênios superar 47 milhões de reais, o total de valores pagos aos favorecidos é de R\$ 8.137.370,07 (Oito milhões

Tabela 9 – Quantidades e valores dos convênios por parlamentar.

<b>Parlamentar</b>	<b>Quant.</b>	<b>%</b>	<b>Valor Total</b>	<b>%</b>
Almeida Lima	2	4,44%	R\$ 6.874.074,61	15,21%
João Daniel	5	11,11%	R\$ 6.679.480,64	14,78%
Maria do Carmo Alves	4	8,89%	R\$ 6.651.661,62	14,72%
Eduardo Amorim	7	15,56%	R\$ 4.051.966,61	8,97%
Fábio Reis	3	6,67%	R\$ 4.022.271,54	8,90%
Jony Marcos	4	8,89%	R\$ 2.329.620,00	5,15%
Virgínio de Carvalho	4	8,89%	R\$ 2.282.900,00	5,05%
André Moura	2	4,44%	R\$ 2.209.010,55	4,89%
Andelson Barreto	2	4,44%	R\$ 2.122.777,59	4,70%
Mendonça Prado	1	2,22%	R\$ 1.976.600,00	4,37%

Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

e cento e trinta e sete mil e trezentos e setenta reais e sete centavos). Deste valor, a empresa "Torre Empreendimentos Rural e Construção LTDA" ficou com o total de R\$ 1.950.404,19 (Um milhão e novecentos e cinquenta mil e quatrocentos e quatro reais e dezenove centavos), o que representa 23,97% do valor total pago. Os três primeiros favorecidos concentram 56,04%, do valor total, equivalente a R\$ 4.559.868,71 (Quatro milhões e quinhentos e cinquenta e nove mil e oitocentos e sessenta e oito reais e setenta e um centavos). Apenas um favorecido apareceu em mais de um convênio, a empresa "AMT Projetos e Serviços LTDA".

Tabela 10 – Quantidades e valores dos convênios por favorecido.

<b>Favorecido</b>	<b>Valor Total</b>	<b>%</b>
Torre Empreendimentos Rural e Construção LTDA	R\$ 1.950.404,19	23,97%
Dicon - Engenharia LTDA - epp	R\$ 1.580.439,23	19,42%
Essencial Transporte & Construções LTDA	R\$ 1.029.025,29	12,65%
Life Comércio e Serviços Eireli	R\$ 706.600,00	8,68%
MGB Execuções Eireli - me	R\$ 544.661,89	6,69%
AMT Projetos e Servicos LTDA	R\$ 495.897,2	6,09%
Labor Med Aparelhagem de Precisão LTDA	R\$ 400.000,00	4,92%
KSS Comércio e Industria de Equip. Médico LTDA	R\$ 170.550,00	2,10%
Medicalway Equip. Médicos LTDA	R\$ 146.000,00	1,79%
Marya Bunita Produções Artísticas LTDA	R\$ 133.697,48	1,64%

Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

## **4.6 Considerações finais do capítulo**

Este capítulo apresentou as etapas e diretrizes para a realização do estudo de caso. Tratou inicialmente sobre o projeto do estudo de caso e seu objetivo principal em responder à questão de pesquisa: "Existem discrepâncias entre os valores dos convênios firmados no estado de Sergipe?"

Na preparação para a coleta de dados, foram tratados dados brutos, criado um projeto de banco de dados e gerado um arquivo com as informações necessárias para a geração das estatísticas no ambiente R.

Após a preparação, execução e coleta de dados, estes foram analisados e avaliados e, ao final, o seu resultado foi consolidado em uma lista que apontou 36 convênios com valores discrepantes, dispersos em 20 diferentes proponentes, atendendo a 8 áreas distintas, com 16 parlamentares responsáveis pelas emendas e 34 favorecidos distintos. O valor total destes convênios é de R\$ 47.961.892,70 (Quarenta e sete milhões e novecentos e sessenta e um mil e oitocentos e noventa e dois reais e setenta centavos).

# 5

## Experimento sobre Associações

Neste capítulo, descreve-se a metodologia utilizada para realização do experimento, bem como os achados oriundos da análise dos resultados experimentais.

### 5.1 Definição e planejamento do experimento

Esta e as próximas seções relatam o processo experimental, de acordo com as diretrizes de (WOHLIN et al., 2012) e (OLIVEIRA; COLACO JÚNIOR, 2018), utilizadas para realização da análise das regras de associação entre parlamentares, áreas, cidades e fornecedores, extraídas dos dados abertos publicados pelo SICONV.

#### 5.1.1 Definição do Objetivo

De acordo com a tabela 11, a distribuição dos convênios ativos propostos por emendas parlamentares individuais, no período de 2008 a 2018, totaliza aproximadamente 45 mil convênios com os respectivos valores na ordem de R\$ 23.3 bilhões, envolvendo convênios com estados, municípios e entidades sem fins lucrativos. Nota-se que cinco estados (na ordem: São Paulo, Minas Gerais, Pernambuco, Bahia, Rio Grande do Sul) concentram 45,47% dos recursos disponibilizados, sendo o estado de São Paulo detentor do maior número de convênios assinados, o que representa 15,96% do total de recursos disponibilizados no período. Enquanto os cinco estados com menores soma de recursos recebidos totalizam pouco mais de R\$ 1,9 bilhão, com uma quantidade de 3,5 mil convênios ativos. Ainda nesta tabela, é mostrada a quantidade de fornecedores, prestadores de serviços, municípios atendidos pelos convênios e a quantidade de parlamentares que propuseram as emendas para execução dos serviços. Apenas 3,3% dos municípios brasileiros não possuem convênios ativos. Em relação à quantidade de parlamentares, este número conta com Deputados e Senadores que não estão mais no exercício do mandato. Estes fazem parte da contagem, pois os convênios que foram criados com suas emendas ainda



estão em execução.

Tabela 11 – Distribuição dos Convênios Ativos por unidade federativa (Valores em R\$ bilhões).

Estado	Sigla	Quant. Conv.	%	Valor Total	%	Quant. Forn.	Quant. Parl.	Quant. Mun.
São Paulo	SP	6776	14.98	3.71	15.96	3092	178	614
Pernambuco	PE	1628	3.60	2.14	9.22	677	45	177
Minas Gerais	MG	5072	11.22	1.90	8.16	1803	95	804
Bahia	BA	3090	6.83	1.57	6.76	902	67	409
Rio Grande do Sul	RS	4121	9.11	1.25	5.37	1800	57	488
Paraná	PR	3009	6.65	1.14	4.89	1386	56	388
Rio de Janeiro	RJ	1360	3.01	1.04	4.49	1112	65	92
Ceara	CE	1499	3.31	0.93	4.00	627	40	180
Maranhão	MA	1271	2.81	0.80	3.46	334	40	210
Goiás	GO	1832	4.05	0.77	3.31	500	35	241
Santa Catarina	SC	1983	4.39	0.65	2.79	840	32	286
Mato Grosso do Sul	MS	801	1.77	0.61	2.61	204	21	78
Piauí	PI	1298	2.87	0.60	2.57	315	23	221
Para	PA	903	2.00	0.57	2.47	278	30	132
Paraíba	PB	1321	2.92	0.55	2.38	325	25	218
Alagoas	AL	659	1.46	0.53	2.28	391	21	97
Espirito Santo	ES	1272	2.81	0.51	2.19	430	22	77
Tocantins	TO	965	2.13	0.48	2.07	483	21	138
Roraima	RR	384	0.85	0.48	2.08	56	22	15
Rondônia	RO	935	2.07	0.46	1.96	231	20	52
Rio Grande do Norte	RN	1281	2.83	0.45	1.95	495	26	165
Acre	AC	910	2.01	0.44	1.91	500	20	22
Mato Grosso	MT	736	1.63	0.42	1.79	227	21	130
Sergipe	SE	850	1.88	0.37	1.60	254	19	74
Amazonas	AM	524	1.16	0.35	1.49	77	14	62
Amapá	AP	509	1.13	0.34	1.44	126	16	16
Distrito Federal	DF	230	0.51	0.19	0.80	247	19	1
Total		45219	100%	23.26	100%	17712	1050	5387

Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

Esses convênios são distribuídos entre os mais variados concedentes, como mostrado na tabela 12, onde são listados os 10 principais concedentes. Estes somam 89,18% dos recursos distribuídos e 88,53% dos convênios assinados no período de 2008 a 2018, demonstrando uma grande concentração perante as dezenas de órgãos que poderiam utilizar este mecanismo de repasse. Entretanto, dentre estes, existem aqueles de abrangência nacional, responsáveis por políticas públicas descentralizadas, a exemplo do Programa Bolsa Família, fundamentando a

importância da distribuição de recursos por meio de convênios, que, normalmente, tornam-se projetos em diversas áreas com forte apelo popular: educação, saúde, transportes, saneamento, esportes e cultura.

Tabela 12 – Total de Convênios e Valores dos principais Ministérios repassadores de 2008 a 2018 (R\$ bilhões)

Órgão	Quant.	%	Valor	%
MINISTÉRIO DAS CIDADES	13536	29.93	6.40	27.51
MINISTÉRIO DA SAÚDE	9350	20.68	6.15	26.44
MINISTÉRIO DO TURISMO	3736	8.26	1.85	7.96
MINISTÉRIO DO ESPORTE	4087	9.04	1.74	7.47
MINISTÉRIO DA AGRIC. PEC E ABAST.	4346	9.61	1.38	5.93
MINISTÉRIO DA DEFESA	1877	4.15	1.07	4.61
FUNDAÇÃO NACIONAL DE SAÚDE	1149	2.54	0.58	2.51
COMP. DE DESENV. V. DO S. FRAN.	1060	2.34	0.53	2.28
MINISTÉRIO DO TRAB. E EMP.	55	0.12	0.52	2.26
MINISTÉRIO DA INTEG. NAC.	842	1.86	0.51	2.21
OUTROS	5181	11,44	2,52	10,8
Total	45219	100%	23,25	100%

Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

Como exposto anteriormente nas tabelas 12 e 11, há um grande número de convênios ativos com valores em reais expressivos. Praticamente todos os municípios brasileiros possuem convênios ativos e um grande número de parlamentares está gerando emendas constitucionais para criar os convênios. Assim, o objetivo deste experimento é analisar as possíveis associações entre os parlamentares, fornecedores ou prestadores de serviços, áreas e os municípios que recebem esses convênios.

Para atingi-lo, conduziu-se um experimento em ambiente controlado, no qual foram verificadas Regras de Associação, utilizando o algoritmo *Apriori*, para determinação das combinações de itens que ocorrem com determinada frequência, bem como para a geração das medidas de interesse que, estatisticamente, puderam servir de base para medir a força das regras de associação.

Segue a formalização do objetivo, segundo o modelo GQM (Goal Question Metric) apresentado em (BASILI et al., 2014):

- **Analisar** convênios de repasses federais
- **com o propósito** de avaliá-los
- **com respeito** às regras de associação entre parlamentares, empresas, cidades e áreas
- **do ponto de vista de** analistas de dados e cidadãos
- **no contexto de** dados abertos do SICONV.

### 5.1.2 Planejamento

**Seleção de Contexto.** O experimento foi *in vitro* e foram utilizados os dados dos arquivos referentes a: Convênios (160.431 registros), Propostas Cadastradas (749.049 registros), Emendas Parlamentares (105.884 registros) e Pagamentos de Favorecidos (3.253.338 registros) para gerar as regras de associação.

**Variáveis independentes.** Conjunto de registros compilados e disponibilizados em arquivos pelo site localizado em (SICONV, 2016), contendo os dados dos convênios, emendas e fornecedores.

**Variáveis dependentes.** Medidas de interesse objetivas para auxiliar na identificação das forças das regras de associação geradas: Suporte, Confiança, *Lift*,  $\rho$  (Coeficiente de correlação de Pearson) e *Qui Quadrado* ( $\chi^2$ ), com seu nível de significância (*p-value*).

**Formulação de Hipóteses.** As questões de pesquisa para esse experimento são:

- **Q1:** Há associações entre parlamentares e áreas específicas?
- **Q2:** Há associações entre parlamentares e favorecidos?
- **Q3:** Há associações entre parlamentares e cidades específicas?

Para responder essas questões de pesquisa, analisaremos as regras de associação geradas para: Paralammentares  $\Rightarrow$  Áreas; Parlammentares  $\Rightarrow$  Favorecidos e Parlammentares  $\Rightarrow$  Cidades. Portanto, as seguintes hipóteses serão testadas para o estado de Sergipe:

- **Hipótese 1 (Q1)**

- H0: Os parlamentares são independentes das áreas específicas.
- H1: Os parlamentares são dependentes de áreas específicas.

- **Hipótese 2 (Q3)**

- H0: Os parlamentares são independentes de empresas específicas.
- H1: Os parlamentares são dependentes de empresas específicas.

- **Hipótese 3(Q4)**

- H0: Os parlamentares são independentes de cidades específicas.
- H1: Os parlamentares são dependentes de cidades específicas.

**Seleção de participantes e objetos.** Foram selecionados apenas os convênios ainda em execução, no estado de Sergipe, no período de 2008 a 2018, propostos por emendas parlamentares nas quais o parlamentar atuou individualmente. Os dados analisados foram obtidos no portal de convênios, localizado em ([SICONV, 2016](#)), utilizando arquivos referentes aos convênios, emendas parlamentares e propostas cadastradas.

**Instrumentação.** O processo de instrumentação teve início com a preparação do ambiente para o experimento e o planejamento da coleta de dados. Em seguida, foi desenvolvido um projeto de banco de dados, utilizando a linguagem de manipulação e criação de dados SQL, o qual incluiu um ETL (Extração, Transformação e Carga) para geração de resultados exportados para arquivos CSV (*Comma Separated Values*), contendo os dados necessários para as regras de associação a serem analisadas. Por fim, os arquivos CSV foram exportados para o ambiente do R.

## 5.2 Operação do Experimento

### 5.2.1 Preparação

A fase de preparação consistiu de: obtenção, por meio de *download*, dos arquivos CVS disponíveis no site [www.portal.convenios.gov.br](http://www.portal.convenios.gov.br); criação do banco de dados no PostgreSQL; execução do ETL para ajustar e importar os dados para o banco criado; e geração dos arquivos CSV com as informações dos conjuntos frequentes, em formato transacional, necessárias para geração das Regras de Associação.

Para cada questão de pesquisa, foi gerado um arquivo CSV, utilizando comandos SQL, com os conjuntos frequentes com dois atributos. Para a questão Q1 (Parlamentares  $\Rightarrow$  Áreas), o arquivo gerado utiliza os atributos "nome do parlamentar" e o "nome do órgão". Já para a questão Q2 (Parlamentares  $\Rightarrow$  Favorecidos), o arquivo gerado utiliza os atributos "nome do parlamentar" e

o "nome do fornecedor". E, por fim, para a Q3 (Parlamentares  $\Rightarrow$  Cidade), o arquivo gerado utiliza os atributos "nome do parlamentar" e o "nome do município". Os comandos SQL utilizados para gerar os arquivos CSVs estão listados no Apêndice C.

### 5.2.2 Execução

Ao final das etapas anteriores, deu-se início ao experimento, com a importação dos arquivos CSV, gerados pelos comandos de consultas SQL, no ambiente do R, onde foram geradas as regras de associações e as medidas de interesse objetivas, utilizando o algoritmo *Apriori* do pacote *arule*.

### 5.2.3 Coleta de Dados

Para cada associação a ser estudada, foi gerado um conjunto de regras com os respectivos valores para as medidas de interesse. Por se tratar de mineração em uma base de dados real, o número de regras gerados foi relativamente alto, como corroborado pelos ensaios apresentados em (ZHENG; KOHAVI; MASON, 2001). Além disso, grande parte destes resultados minerados costuma ser composta por regras óbvias, redundantes ou, até mesmo, contraditórias. Para filtrar as regras interessantes para o estudo, o suporte mínimo foi estabelecido em 3% (aproximadamente 19 transações) e a confiança mínima em 60% (GONSALVES, 2004). Para cada regra de associação, foram gerados, também, os valores para o coeficiente de correlação de *pearson* ( $r$ ), o qui quadrado ( $\chi^2$ ) e o seu nível de significância ( $p$ -value).

### 5.2.4 Validação dos Dados

Para assegurar a análise, interpretação e validação dos resultados, foi aplicado o teste de significância da hipótese, utilizando o teste do qui quadrado ( $\chi^2$ ). Além disso, também foi calculado o coeficiente de correlação de *Pearson* ( $r$ ).

## 5.3 Resultados

Para responder às questões experimentais, os arquivos CSV, contendo as regras em formato transacional, foram gerados pelo projeto de banco de dados e analisados no ambiente R (R Core Team, 2018). Para cada regra a ser estudada, foi gerado um arquivo. Na tabela 13, é exibido um quantitativo de transações para cada uma das regras a ser estudada. E, na tabela 14, são exibidas as regras de associação encontradas com os parâmetros definidos anteriormente.

### 5.3.1 Análise e Interpretação

Na tabela 13, observou-se o número de transações analisadas para cada associação a ser estudada, bem como os itens mais frequentes e a quantidade de regras de associação encontradas.

Tabela 13 – Quantitativos de transações analisadas para cada regra.

Associações	Transações	Frequentes	Regras
Parlamentar e Área	654	Min. das Cidades: 305 Min. do Turismo: 110 Min. do Esporte: 84 Eduardo Amorim: 83 Antônio C. Valadares: 77	2
Parlamentar e Favorecido	515	Antônio C. Valadares: 86 Eduardo Amorim: 60 Valadares Filho: 50 André Moura: 36 Fábio Mitidieri: 36	0
Parlamentar e Cidade	654	Eduardo Amorim: 84 Antônio C. Valadares: 77 André Moura: 69 Laercio Oliveira: 60 Fábio Reis: 52	0

Fonte: Elaborada pelo autor. Informações extraídas dos dados públicos do SICONV

Observa-se que apenas duas regras de associação foram encontradas, para as associações entre Parlamentares e Áreas. Estas regras são detalhadas na tabela 14, juntamente com os valores de suas medidas de interesse: suporte (*supp*), confiança (*conf*), *lift*, quantidade (*count*), coeficiente de correlação de pearson (*r*), qui quadrado ( $\chi^2$ ) e o seu *p-value*.

Entretanto, é preciso que analisemos estas tabelas sob a luz da estatística, para encontrarmos evidências que corroborem, ou não, com as regras de associações encontradas. Para tal, adotamos um nível de confiança de 95% ( $\alpha = 0,05$ ) para todo o experimento.

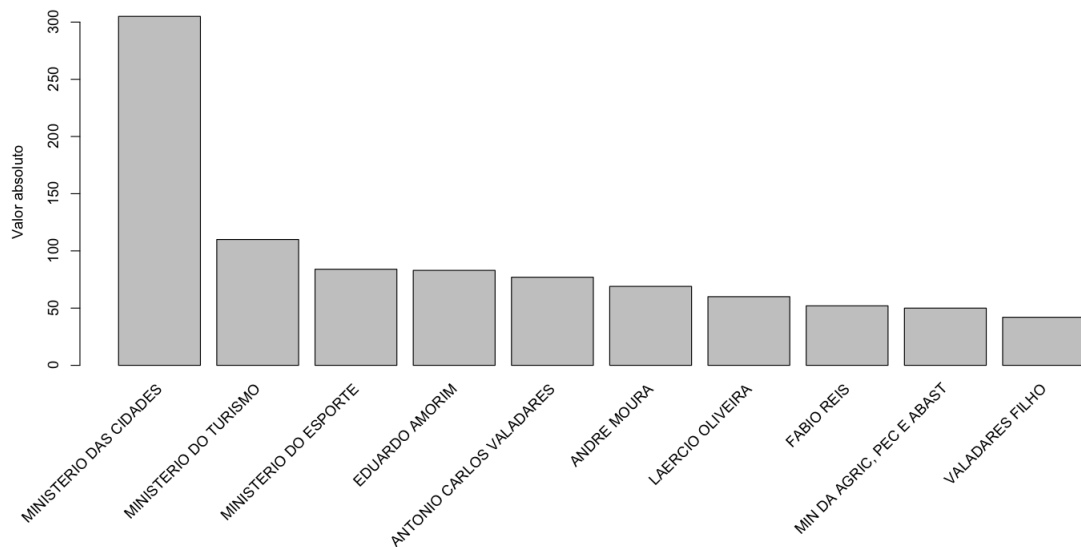
Para melhorar a compreensão, separamos a análise pelas associações, facilitando a visualização das regras de associação encontradas.

### 5.3.1.1 Associação entre Parlamentares e Áreas

Iniciaremos a análise pela quantidade de transações a serem estudadas. No total, foram geradas 654 transações, associando os parlamentares a áreas específicas, como demonstrado na tabela 13. A figura 19 exibe a distribuição dos 10 itens mais frequentes. Nota-se que a área representada pelo Ministério das Cidades e o parlamentar Eduardo Amorim possuem o maior número de convênios ativos.

As regras de associação geradas estão listadas na tabela 14 e na figura 20. Observa-se que apenas duas regras foram geradas. A primeira regra, "Adelson Barreto  $\Rightarrow$  Min. das Cidades", possui um suporte no valor de 4,6%, o que indica uma ocorrência de 30 transações, diante da amostra estudada. Já a confiança, com o valor de 0,91, confere-nos que mais de 90% das

Figura 19 – Frequência das Áreas ou Parlamentares.



Fonte: Elaborada pelo autor. Informações extraídas dos dados públicos do SICONV

transações do parlamentar Adelson Barreto são voltadas para o Min. das Cidades (Antecedente e Consequente). Na segunda regra, "Laercio Oliveira  $\Rightarrow$  Min. das Cidades", observamos um valor um pouco maior no suporte, 6,3%, conferindo uma ocorrência de 41 transações. No entanto, o valor da confiança indica que 68% das transações de Laercio Oliveira estão associadas ao Ministério das Cidades.

Tabela 14 – Regras encontradas para associações entre Parlamentares e Áreas.

Regras	<i>supp</i>	<i>conf</i>	<i>lift</i>	<i>count</i>	<i>r</i>	$\chi^2$	<i>p-value do <math>\chi^2</math></i>
Adelson Barreto $\Rightarrow$ Min. das Cidades	0.04587	0.9091	1.9493	30	0.2046	27.3721	0.0000002
Laercio Oliveira $\Rightarrow$ Min. das Cidades	0.0627	0.6833	1.4652	41	0.1382	12.4963	0.000408

Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

Ambas as regras possuem um valor para o *p-value*, abaixo do nível de significância adotado pelo experimento. Assim, a hipótese  $H_0$  foi rejeitada, indicando uma dependência entre o parlamentar Adelson Barreto e a área representada pelo Ministério das Cidades. O mesmo ocorre com o parlamentar Laercio Oliveira e a área representada pelo Ministério das Cidades.

Figura 20 – Regras de associação do site: Transparência Traduzida.

<b>Associações</b> As associações apresentadas são aquelas com significância estatística. Os parlamentares não listados possuem convênios mais distribuídos e não possuem uma ligação forte com uma área específica. No estado de Sergipe, foi(ram) encontrada(s) a(s) seguinte(s):
ADELSON BARRETO » MINISTERIO DAS CIDADES  <b>ADELSON BARRETO</b> está associado(a) ao(à) <b>MINISTERIO DAS CIDADES</b> . Convênios liderados por <b>ADELSON BARRETO</b> para esta área aparecem em <b>4.6%</b> das ocorrências de convênios. O que significa dizer que são <b>30</b> convênios em um total aproximado de <b>652</b> convênios ativos. A confiança desta associação é de <b>90.9%</b> , ou seja, <b>90.9%</b> dos convênios do(a) parlamentar <b>ADELSON BARRETO</b> são voltados para o <b>MINISTERIO DAS CIDADES</b> . Esta dependência possui uma significância estatística de <b>99.99%</b> .
LAERCIO OLIVEIRA » MINISTERIO DAS CIDADES  <b>LAERCIO OLIVEIRA</b> está associado(a) ao(à) <b>MINISTERIO DAS CIDADES</b> . Convênios liderados por <b>LAERCIO OLIVEIRA</b> para esta área aparecem em <b>6.3%</b> das ocorrências de convênios. O que significa dizer que são <b>41</b> convênios em um total aproximado de <b>651</b> convênios ativos. A confiança desta associação é de <b>68.3%</b> , ou seja, <b>68.3%</b> dos convênios do(a) parlamentar <b>LAERCIO OLIVEIRA</b> são voltados para o <b>MINISTERIO DAS CIDADES</b> . Esta dependência possui uma significância estatística de <b>99.95%</b> .

Fonte: [www.transparenciatraduzida.com.br](http://www.transparenciatraduzida.com.br)

O segundo cálculo realizado foi o do coeficiente de correlação de *Pearson* ( $r$ ). Em ambas as regras, o valor de  $r$  foi maior que 0 (zero), demonstrando uma correlação positiva baixa, ou seja, quando a frequência do "Antecedente" aumentar, a frequência do "Consequente" aumenta, na proporção indicada por  $r$ .

### 5.3.1.2 Associação entre Parlamentares e Favorecidos

A análise da quantidade de transações geradas pela associação entre Parlamentares e Favorecidos totaliza 515 transações, como demonstrado na tabela 13. A figura 21 exibe a distribuição dos 10 itens mais frequentes. Nota-se que os parlamentares mais frequentes são: Antônio Carlo Valadares, seguido por Eduardo Amorim e Valadares Filho. Nesta figura, não se observa nenhum favorecido (fornecedor) destacado, mostrando uma boa distribuição de convênios entre os diferentes favorecidos.

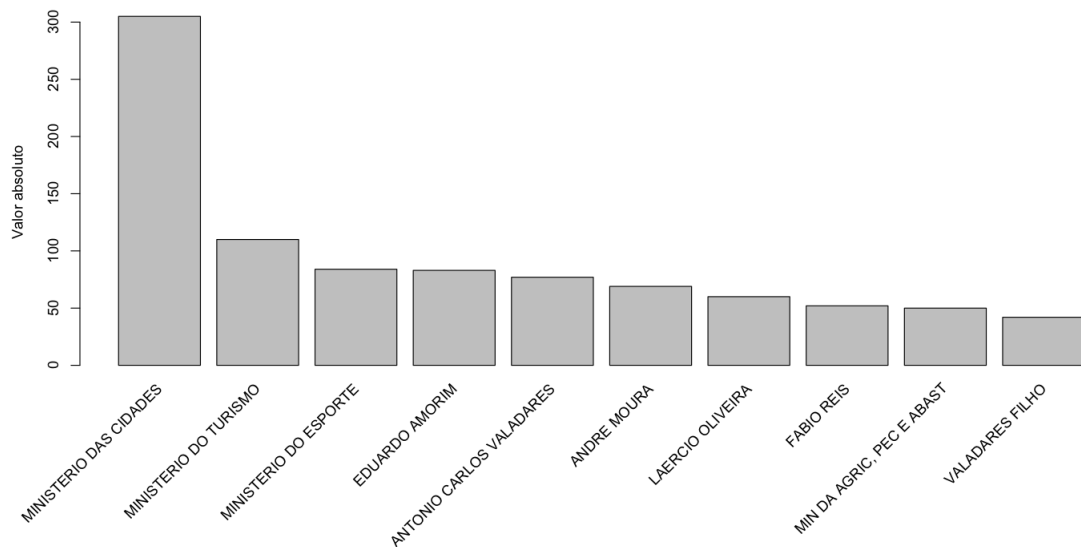
Não foram encontradas regras de associação entre Parlamentares e Favorecidos, o que nos leva a não refutar a  $H_0$ . Ou seja, ficou evidenciado que não existem associações entre Parlamentares e Favorecidos, nos convênios ainda em execução, propostos por emendas parlamentares individuais.

### 5.3.1.3 Associação entre Parlamentares e Cidades

A análise da quantidade de transações geradas pelas associações entre Parlamentares e Cidades totaliza 654 transações, como demonstrado na tabela 13. A figura 22 exibe a distribuição dos 10 itens mais frequentes. Nota-se que o parlamentar mais frequente é Eduardo Amorim, seguido por Antônio Carlo Valadares e André Moura. Os três parlamentares com valores aproximados. Nesta figura, não se observa nenhuma cidade (localidade) destacada, mostrando um equilíbrio na distribuição de convênios entre as diferentes cidades.



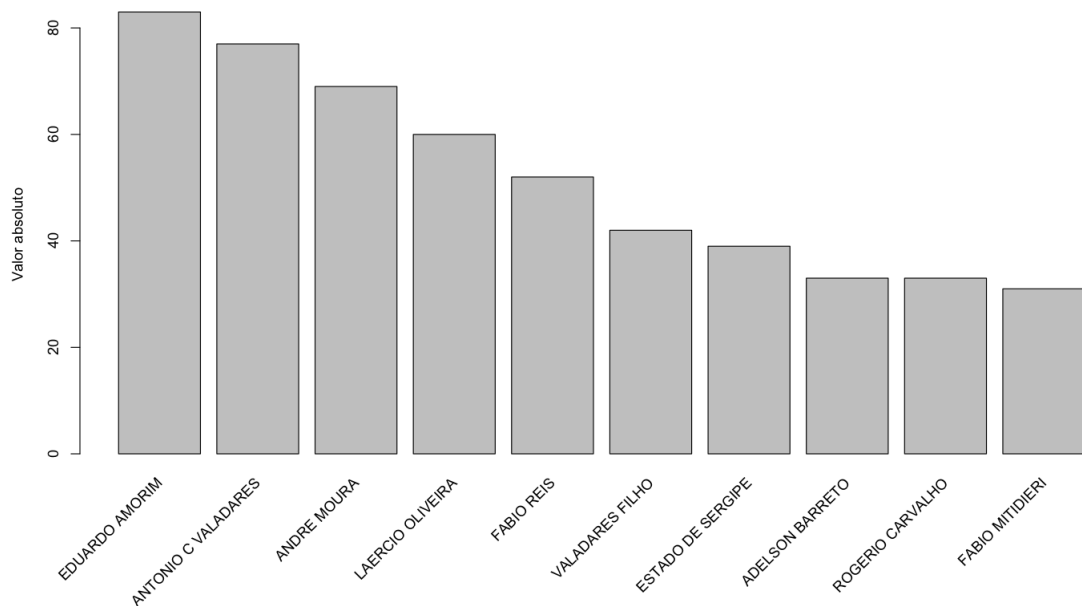
Figura 21 – Frequência dos Parlamentares e Favorecidos.



Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

Figura 22 – Frequência dos Parlamentares e Cidades.



Fonte: Elaborada pelo autor (2018).

Nota: Informações extraídas dos dados públicos do SICONV.

Não foram encontradas regras de associação entre Parlamentares e Cidades, o que nos leva a não refutar a  $H_0$ . Ou seja, ficou evidenciado que não existem associações entre Parlamentares e Cidades, nos convênios ainda em execução, propostos por emendas parlamentares individuais.

### 5.3.2 Ameaças à Validade

Embora os resultados obtidos por meio do experimento tenham se mostrado satisfatórios, o mesmo apresenta ameaças a sua validade que precisam ser consideradas:

#### 5.3.2.1 Ameaças à validade interna

Considerando que os dados foram obtidos, por meio de *download*, tratados e analisados pelos autores, há uma forte ameaça à validade interna. No entanto, para mitigar possíveis erros, todos os artefatos de *software* construídos para o tratamento dos dados, e os resultados por eles gerados, tiveram sua construção e seus resultados revisados por mais de um pesquisador.

#### 5.3.2.2 Ameaças à validade externa

Considerando a avaliação realizada no experimento, os resultados obtidos dizem respeito apenas aos convênios em execução propostos por emendas parlamentares individuais no estado de Sergipe, causando uma redução na amostra estudada. No entanto, essa escolha nos permitiu analisar as atividades do parlamentar individualmente.

## 5.4 Considerações Finais

Neste capítulo, realizou-se o experimento responsável por analisar as possíveis associações entre: "Parlamentares e Áreas", "Parlamentares e Cidades" e "Parlamentares e Favorecidos". Por meio do estudo de Regras de Associação, podemos observar que cada uma das associações estudadas apresentou características próprias, reforçando a necessidade de estudá-las individualmente.

Após a execução do experimento, destacaram-se apenas duas regras de associação com valores estatisticamente relevantes. Ambas as regras foram sobre "Parlamentares e Áreas". Foram elas: Adelson Barreto com o Min. das Cidades e Laercio Oliveira, também, com o Min. das Cidades. Não foi localizada nenhuma regra com valores estatisticamente relevantes para as regras de associação "Parlamentares e Cidades" e "Parlamentares e Favorecidos".

Vale ressaltar que a análise individual das regras de associação é uma forma de dar transparência aos dados brutos do SICONv.

# 6

## Conclusão

As recentes políticas de disponibilização de dados públicos pelo governo federal devem estimular a participação da população na fiscalização da transparência da gestão. Esta participação pode ser potencializada e estimulada, desde que os dados fornecidos pelo governo sejam estruturados e dispostos de uma forma entendível, ou efetivamente transparente, para o cidadão. Quando isto não ocorre, pesquisas sobre dados abertos podem contribuir com a estruturação, autenticação, interpretação, organização, tradução, descoberta de padrões e disponibilização de informação útil para tomada de decisão e para o controle social exercido pela população.

### 6.1 Contribuições

A principal contribuição deste estudo consiste na condução de duas avaliações: um estudo de caso para verificação de valores aberrantes (*outliers*) entre os convênios ainda em execução e propostos por emendas parlamentares individuais; e a execução de um processo experimental para analisar as regras de associação entre "parlamentares e áreas", "parlamentares e municípios" e "parlamentares e favorecidos", no estado de Sergipe.

Em resumo, as principais contribuições deste trabalho são:

- Mapeamento Sistemático utilizado para identificar e sistematizar as abordagens e técnicas computacionais de análise inteligente de dados que têm sido utilizadas para dar transparência, especificamente, aos repasses federais, ou para favorecer a Governança Pública em geral. Ressalta-se que os resultados foram publicados no 15<sup>th</sup> CONTECSI USP - *International Conference on Information Systems and Technology Management* - ISSN 1041-2448. Realizado na USP, no período de 23 a 25 de maio, 2018, São Paulo, Brasil;
- Criação de ETL e um projeto de banco de dados para automatizar o tratamento dos dados

brutos disponibilizados pelo SICONV;

- Estudo de caso para análise dos valores aberrantes (*outliers*) dos convênios ainda em execução e propostos por emendas parlamentares individuais;
- Avaliação experimental para análise de regras de associação entre "parlamentares e áreas", "parlamentares e municípios" e "parlamentares e favorecidos", no estado de Sergipe;
- Aplicação Web denominada de Transparência Traduzida,<sup>1</sup> com acesso público, que torna mais fácil a consulta e visualização dos resultados.

Como consequência destas contribuições, conseguimos responder às questões de pesquisa elaboradas no início do trabalho:

- Q1: Existem discrepâncias entre os valores dos convênios firmados? Sim. O estudo de caso mostrou um total de 36 convênios com valores aberrantes (*outliers*) no estado de Sergipe, atendendo a 20 diferentes proponentes, 8 áreas distintas, 16 parlamentares, 34 favorecidos e totalizado um valor de R\$ 47.961.892,70 (Quarenta e sete milhões e novecentos e sessenta e um mil e oitocentos e noventa e dois reais e setenta centavos).
- Q2: Há associações de parlamentares com áreas específicas? Sim. Foram encontradas duas regras de associação entre os parlamentares Andelson Barreto e Laercio Oliveira, com a área representada pelo Ministério das Cidades.
- Q3: Há associações entre parlamentares e favorecidos? Não. Nenhuma regra de associação foi encontrada entre Parlamentares e Favorecidos, nos convênios ainda em execução e propostos por emendas parlamentares individuais, no estado de Sergipe.
- Q4: Há associações entre parlamentares e cidades específicas? Não. Nenhuma regra de associação foi encontrada entre Parlamentares e Cidades, nos convênios ainda em execução e propostos por emendas parlamentares individuais no estado de Sergipe.

## 6.2 Limitações

O estudo realizado apresentou limitações importantes quanto ao tratamento dos dados e na definição da amostra.

Os dados utilizados nesta pesquisa foram obtidos, por meio de *download*, tratados e analisados pelos autores. Sendo assim, todos os artefatos de *software* construídos para o tratamento dos dados e os resultados por eles gerados foram revisados por mais de um pesquisador, com o intuito de mitigar possíveis erros. No entanto, esta tarefa de revisão, realizada pelos pesquisadores, introduz um fator subjetivo, visto que depende de uma observação cuidadosa do

---

<sup>1</sup> [www.transparenciatraduzida.com.br](http://www.transparenciatraduzida.com.br)

pesquisador para o perfeito enquadramento das informações. Com isso, interpretações errôneas podem gerar análises inconsistentes.

Em relação à definição da amostra, deve-se considerar que os dados selecionados para a pesquisa foram restringidos aos convênios em execução e propostos por emendas parlamentares individuais no estado de Sergipe. Com essas restrições, um grande número de convênios propostos por bancadas parlamentares não entraram nas análises realizadas pela pesquisa.

## 6.3 Perspectivas

A grande quantidade de informações brutas disponibilizadas, no portal dos convênios, ([SICONV, 2016](#)) possibilita-nos criar inúmeras avaliações nestes dados. Além da pesquisa desenvolvida neste trabalho, outros possíveis desdobramentos são:

- Elaborar um modelo padrão de descrição dos dados de projetos realizados por meio de convênios federais;
- Os recursos distribuídos, pelas emendas parlamentares, são proporcionais entre as diferentes áreas atendidas?
- Os recursos distribuídos, pelas emendas parlamentares, são proporcionais à população?

## 6.4 Considerações Finais

Este trabalho apresentou os resultados da análise de *outliers*, nos valores dos convênios ainda em execução, bem como regras de associação entre parlamentares, áreas, municípios ou empresas, por meio de um processo experimental. Por fim, relatou-se as contribuições obtidas, bem como as perspectivas de pesquisas futuras.

Os dados aqui apresentados são apenas uma tradução dos dados brutos brasileiros disponibilizados para *download*, no site do portal de convênios. **Esta pesquisa buscou apenas identificar padrões e fazer a efetiva tradução das informações, o que não implica que estão sendo feitos quaisquer tipos de julgamentos com relação aos nomes citados ou que há algum descumprimento perante às leis vigentes. São apenas evidências, cabendo às autoridades competentes a análise e julgamento de qualquer descumprimento da lei.**

# Referências

AGRAWAL, R.; IMIELIŃSKI, T.; SWAMI, A. Mining association rules between sets of items in large databases. In: ACM. *Acm sigmod record*. [S.l.], 1993. v. 22, n. 2, p. 207–216. Citado na página 30.

ALBUQUERQUE, D. J. S. et al. Implementing e-government processes distribution with transparency using multi-agent systems. *iSys-Revista Brasileira de Sistemas de Informação*, v. 9, n. 1, p. 118–138, 2016. Citado na página 24.

BASIL, V. et al. *Aligning Organizations Through Measurement: The GQM+Strategies Approach*. Springer International Publishing, 2014. (The Fraunhofer IESE Series on Software and Systems Engineering). ISBN 9783319050478. Disponível em: <<https://books.google.com.br/books?id=oJHIAwAAQBAJ>>. Citado na página 57.

BRASIL. *Constituição da República Federativa do Brasil*. Brasília, DF, 1988. Acessado em 02/03/2017. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/Constituicao/Constituicao.htm](http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm)>. Citado na página 17.

BRASIL. Palácio do Planalto. *Medida Provisória nº 2.143-31, de 2 de abril de 2001*. Brasília, 2001. Acessado em 06/12/2016. Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/mpv/Antigas\\_2001/2143-31.htm](http://www.planalto.gov.br/ccivil_03/mpv/Antigas_2001/2143-31.htm)>. Citado na página 15.

BRASIL. Palácio do Planalto. *LEI Nº 10.683, DE 28 DE MAIO DE 2003*. Brasília: Poder Executivo, 2003. Acessada em 06/12/2016. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/leis/2003/110.683.htm](https://www.planalto.gov.br/ccivil_03/leis/2003/110.683.htm)>. Citado na página 15.

BRASIL. Palácio do Planalto. *DECRETO Nº 5.683, DE 24 DE JANEIRO DE 2006*. Brasília: Poder Executivo, 2006. Acessada em 06/12/2016. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/\\_Ato2004-2006/2006/Decreto/D5683.htm](https://www.planalto.gov.br/ccivil_03/_Ato2004-2006/2006/Decreto/D5683.htm)>. Citado na página 15.

BRASIL. Palácio do Planalto. *DECRETO Nº 8.109, DE 17 DE SETEMBRO DE 2013*. Brasília, 2013. Acessada em 06/12/2016. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/\\_Ato2011-2014/2013/Decreto/D8109.htm#art6](https://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2013/Decreto/D8109.htm#art6)>. Citado na página 15.

BRASIL. Palácio do Planalto. Brasília, 2016. Acessada em 06/12/2016. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/\\_Ato2015-2018/2016/Decreto/D8910.htm#art11](https://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2016/Decreto/D8910.htm#art11)>. Citado na página 16.

BRASIL, T. *CONVÊNIOS E OUTROS REPASSES - 6ª EDIÇÃO*. 2016. Acessado em 05/04/2017. Disponível em: <<https://portal.tcu.gov.br/biblioteca-digital/convenios-e-outros-repasses-6-edicao.htm>>. Citado na página 25.

BRASIL, T. *Transparência Brasil, Quem Somos*. 2016. Disponível em: <http://www.transparencia.org.br/>. Acessado em 01/12/2016. Citado na página 15.

CONFORTI, R.; ROSA, M. L.; HOFSTEDE, A. H. ter. Noise filtering of process execution logs based on outliers detection. 2015. Citado na página 28.

- CRUZ, C. F. et al. Transparência da gestão pública municipal: um estudo a partir dos portais eletrônicos dos maiores municípios brasileiros. *Revista de Administração Pública*, SciELO Brasil, v. 46, n. 1, p. 153–176, 2012. Citado na página 24.
- DESSÌ, N. et al. Increasing open government data transparency with spatial dimension. In: *IEEE. Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), 2016 IEEE 25th International Conference on*. [S.l.], 2016. p. 247–249. Citado na página 21.
- FAYYAD, U.; PIATETSKY-SHAPIO, G.; SMYTH, P. From data mining to knowledge discovery in databases. *AI magazine*, v. 17, n. 3, p. 37, 1996. Citado na página 27.
- FREITAS, R. K. V. de; DACORSO, A. L. R. Inovação aberta na gestão pública: análise do plano de ação brasileiro para a open government partnership. *Revista de Administração Pública*, v. 48, n. 4, p. 869–888, 2014. Citado na página 15.
- GONSALVES, E. C. Regras de associações e suas medidas de interesse objetivas e subjetivas. *INFOCOMP*, v. 4, n. 1, p. 26–35, 2004. Citado 4 vezes nas páginas 30, 31, 32 e 59.
- HAN, J.; PEI, J.; KAMBER, M. *Data mining: concepts and techniques*. [S.l.]: Elsevier, 2011. Citado 2 vezes nas páginas 28 e 31.
- HARRISON, T. M. et al. Open government and e-government: Democratic challenges from a public value perspective. *Information Polity*, IOS Press, v. 17, n. 2, p. 83–97, 2012. Citado 2 vezes nas páginas 15 e 24.
- INTERNATIONAL, T. *Transparency International, WHO WE ARE - Our History*. 2016. Disponível em <http://www.transparency.org/whoweare/history>. Acessado em 01/12/2016. Citado na página 15.
- JANSSEN, M.; CHARALABIDIS, Y.; ZUIDERWIJK, A. Benefits, adoption barriers and myths of open data and open government. *Information systems management*, Taylor & Francis, v. 29, n. 4, p. 258–268, 2012. Citado na página 15.
- JURISTO, N.; MORENO, A. M. *Basics of software engineering experimentation*. [S.l.]: Springer Science & Business Media, 2013. Citado na página 23.
- KITCHENHAM, B. Procedures for performing systematic reviews. *Keele, UK, Keele University*, v. 33, n. 2004, p. 1–26, 2004. Citado na página 34.
- KUMAR, S.; JOSHI, N. et al. Rule power factor: a new interest measure in associative classification. *Procedia Computer Science*, Elsevier, v. 93, p. 12–18, 2016. Citado 2 vezes nas páginas 30 e 31.
- LAURIKKALA, J. et al. Informal identification of outliers in medical data. In: *Fifth International Workshop on Intelligent Data Analysis in Medicine and Pharmacology*. [S.l.: s.n.], 2000. v. 1, p. 20–24. Citado na página 28.
- LENCA, P. et al. Association rule interestingness measures: Experimental and theoretical studies. In: *Quality Measures in Data Mining*. [S.l.]: Springer, 2007. p. 51–76. Citado na página 32.
- MATHEUS, R.; JANSSEN, M. Exploitation and exploration strategies to create data transparency in the public sector. In: *ACM. Proceedings of the 9th international conference on theory and practice of electronic governance*. [S.l.], 2016. p. 13–16. Citado 3 vezes nas páginas 14, 15 e 25.

- OLIVEIRA, A. F. d. et al. *Convênios e contratos de repasse: análise da prestação de contas pelo Controle Interno*. [S.l.]: Brasília, 2011. 13–23 p. Citado na página 17.
- OLIVEIRA, R. A. N. de; COLACO JÚNIOR, M. Summary report of experimental analysis of stemming algorithms applied to judicial jurisprudence. In: *Information Technology-New Generations*. [S.l.]: Springer, 2018. p. 935–937. Citado na página 54.
- PETERSEN, K. et al. Systematic mapping studies in software engineering. In: *EASE*. [S.l.: s.n.], 2008. v. 8, p. 68–77. Citado na página 34.
- R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2018. Disponível em: <<http://www.R-project.org/>>. Citado na página 59.
- ROMÃO, W. et al. Extração de regras de associação em c&t: O algoritmo apriori. *XIX Encontro Nacional em Engenharia de Produção*, v. 34, p. 37–39, 1999. Citado na página 30.
- SETTLEMENTS, B. for I. *Basel Committee on Banking Supervision. Bank for International Settlements*. 2006. Disponível em: <<http://www.bis.org/publ/bcbs128.pdf>>. Citado na página 15.
- SICONV, P. dos C. *Portal de Convênios*. 2016. Acessado em 19/06/2016. Disponível em: <[portal.convenios.gov.br/](http://portal.convenios.gov.br/)>. Citado 7 vezes nas páginas 16, 20, 42, 43, 57, 58 e 67.
- SILVA, C. V. S.; RALHA, C. G. Detecção de cartéis em licitações públicas com agentes de mineração de dados. Faculdade Cenecista de Campo Largo-Paraná-Brasil, 2011. Citado na página 27.
- SILVA, C. V. S.; ROCHA, H. A. da. Técnicas de mineração de dados como apoio às auditorias governamentais. *CGU*, p. 29, 2010. Citado 2 vezes nas páginas 16 e 27.
- TAN, P.-N.; KUMAR, V.; SRIVASTAVA, J. Selecting the right interestingness measure for association patterns. In: *ACM. Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*. [S.l.], 2002. p. 32–41. Citado na página 32.
- TURNES, P. B.; ERNST, R. A framework for transparency in international trade. *Investigaciones Europeas de Dirección y Economía de la Empresa*, Elsevier, v. 21, n. 1, p. 1–8, 2015. Citado na página 15.
- URBINATI, N. *Representative democracy: principles and genealogy*. [S.l.]: University of Chicago Press, 2006. Citado na página 14.
- WOHLIN, C. et al. *Experimentation in software engineering*. [S.l.]: Springer Science & Business Media, 2012. Citado 3 vezes nas páginas 17, 22 e 54.
- WU, J. et al. Computing exact permutation p-values for association rules. *Information Sciences*, Elsevier, v. 346, p. 146–162, 2016. Citado na página 31.
- ZHENG, Z.; KOHAVI, R.; MASON, L. Real world performance of association rule algorithms. In: *ACM. Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*. [S.l.], 2001. p. 401–406. Citado na página 59.



# **Apêndices**

# APÊNDICE A – Script: Criação do base de dados

```

create table programas_disponibilizados
(
COD_ORGAO_SUP_PROGRAMA      int          null,
DESC_ORGAO_SUP_PROGRAMA     varchar(100) null,
ID_PROGRAMA                  int          null,
COD_PROGRAMA                 bigint       null,
NOME_PROGRAMA                varchar(100) null,
SIT_PROGRAMA                 varchar(100) null,
DATA_DISPONIBILIZACAO       date         null,
ANO_DISPONIBILIZACAO        int          null,
DT_PROG_INI_RECEB_PROP      date         null,
DT_PROG_FIM_RECEB_PROP      date         null,
DT_PROG_INI_EMENDA_PAR      date         null,
DT_PROG_FIM_EMENDA_PAR      date         null,
DT_PROG_INI_BENEF_ESP       date         null,
DT_PROG_FIM_BENEF_ESP       date         null,
MODALIDADE_PROGRAMA         varchar(100) null,
NATUREZA_JURIDICA_PROGRAMA  varchar(100) null,
UF_PROGRAMA                  char(2)      null );

create table emendas_parlamentares
(
ID_PROPOSTA                  bigint       null,
QUALIF_PROPONENTE            varchar(100) null,
COD_PROGRAMA_EMENDA         bigint       null,
NR_EMENDA                    bigint       null,
NOME_PARLAMENTAR             varchar(100) null,
BENEFICIARIO_EMENDA         varchar(100) null,
IND_IMPOSITIVO               varchar(100) null,
TIPO_PARLAMENTAR             varchar(100) null,
VALOR_REPASSE_PROPOSTA_EMENDA numeric(17,2) null,
VALOR_REPASSE_EMENDA        numeric(17,2) null );

create table plano_aplicacao_detalhado
(
ID_PROPOSTA                  bigint       null,
SIGLA                        char(2)      null,
MUNICIPIO                    varchar(100) null,
NATUREZA_AQUISICAO           int          null,
DESCRICAO_ITEM               text         null,
CEP_ITEM                     varchar(100) null,
ENDERECO_ITEM                varchar(100) null,
TIPO_DESPESA_ITEM            varchar(100) null,
NATUREZA_DESPESA             varchar(100) null,
SIT_ITEM                     varchar(100) null,
QTD_ITEM                     int          null,
VALOR_UNITARIO_ITEM          numeric(15,2) null,
VALOR_TOTAL_ITEM             numeric(15,2) null );

create table desembolsos
(
NR_CONVENIO                  bigint       null,
DT_ULT_DESEMBOLSO           date         null,
QTD_DIAS_SEM_DESEMBOLSO     int          null,
ID_DESEMBOLSO                bigint       null,
DATA_DESEMBOLSO             date         null,
ANO_DESEMBOLSO              int          null,
MES_DESEMBOLSO              int          null,
NR_SIAFI                     varchar(100) null,
VL_DESEMBOLSADO              numeric(15,2) null );

create table programas_propostas
(
ID_PROGRAMA                  bigint       null,
ID_PROPOSTA                  bigint       null );

create table propostas_cadastradas
(
ID_PROPOSTA                  bigint       null,
UF_PROPONENTE                char(2)      null,
MUNIC_PROPONENTE             varchar(100) null,
COD_MUNIC_IBGE               int          null,
COD_ORGAO_SUP                int          null,
DESC_ORGAO_SUP               varchar(100) null,
NATUREZA_JURIDICA            varchar(100) null,
NR_PROPOSTA                  int          null,
DIA_PROP                     int          null,
MES_PROP                     int          null,
ANO_PROP                     int          null,
DIA_PROPOSTA                 int          null,
COD_ORGAO                    int          null,
DESC_ORGAO                   varchar(100) null,
MODALIDADE                   varchar(100) null,
IDENTIF_PROPONENTE           varchar(15) null,
NM_PROPONENTE                 varchar(100) null,
CEP_PROPONENTE                varchar(8) null,
ENDERECO_PROPONENTE          varchar(100) null,
BAIRRO_PROPONENTE            varchar(100) null,
NM_BANCO                     varchar(100) null,
SITUACAO_CONTA                varchar(100) null,
SITUACAO_PROJETO_BASICO       varchar(100) null,
SIT_PROPOSTA                  varchar(100) null,
DIA_INIC_VIGENCIA_PROPOSTA   date         null,
DIA_FIM_VIGENCIA_PROPOSTA    date         null,
OBJETO_PROPOSTA               varchar(200) null,
VL_GLOBAL_PROP                numeric(15,2) null,
VL_REPASSE_PROP               numeric(15,2) null,
VL_CONTRAPARTIDA_PROP         numeric(15,2) null );

```

```

create table convenioassinado
(
  NR_CONVENIO          bigint          null,
  ID_PROPOSTA          bigint          null,
  DIA                  smallint        null,
  MES                  smallint        null,
  ANO                  smallint        null,
  DIA_ASSIN_CONV       date            null,
  SIT_CONVENIO         varchar(100)    null,
  SUBSITUACAO_CONV     varchar(100)    null,
  SITUACAO_PUBLICACAO  varchar(100)    null,
  INSTRUMENTO_ATIVO    varchar(10)     null,
  IND_OPERA_OBTV       varchar(10)     null,
  NR_PROCESSO          varchar(100)    null,
  UG_EMITENTE          varchar(100)    null,
  DIA_PUBL_CONV        date            null,
  DIA_INIC_VIGENC_CONV date            null,
  DIA_FIM_VIGENC_CONV  date            null,
  DIAS_PREST_CONTAS    int             null,
  DIA_LIMITE_PREST_CONTAS date          null,
  SITUACAO_CONTRATACAO varchar(100)    null,
  IND_ASSINADO         varchar( 10)    null,
  QTDE_CONVENIOS       int             null,
  QTD_TA               int             null,
  QTD_PRORROGA         int             null,
  VL_GLOBAL_CONV       numeric(17,2)   null,
  VL_REPASSE_CONV      numeric(17,2)   null,
  VL_CONTRAPARTIDA_CONV numeric(17,2)   null,
  VL_EMPENHADO_CONV    numeric(17,2)   null,
  VL_DESEMBOLSADO_CONV numeric(17,2)   null,
  VL_SALDO_REMAN_TESOURO numeric(17,2)   null,
  VL_SALDO_REMAN_CONVENENTE numeric(17,2)   null,
  VL_RENDIMENTO_APLICACAO numeric(17,2)   null,
  VL_INGRESSO_CONTRAPARTIDA numeric(17,2) null );

create table historico_situacao_propostas
(
  ID_PROPOSTA          bigint          null,
  NR_CONVENIO          bigint          null,
  DIA_HISTORICO_SIT    date            null,
  HISTORICO_SIT        varchar(100)    null,
  DIAS_HISTORICO_SIT   int             null,
  COD_HISTORICO_SIT    int             null );

create table empenhos_realizados
(
  NR_CONVENIO          bigint          null,
  NR_EMPENHO           varchar(100)    null,
  TIPO_NOTA            int             null,
  DESC_TIPO_NOTA       varchar(100)    null,
  DATA_EMISSAO        date            null,
  COD_SITUACAO_EMPENHO int             null,
  DESC_SITUACAO_EMPENHO varchar(100)    null,
  VALOR_EMPENHO        numeric(15,2)   null );

create table ingresso_contrapartida
(
  NR_CONVENIO          bigint          null,
  DT_INGRESSO_CONTRAPARTIDA date        null,
  VL_INGRESSO_CONTRAPARTIDA numeric(15,2) null );

create table pagamentos_favorecidos
(
  NR_MOV_FIN           int             null,
  NR_CONVENIO          bigint          null,
  IDENTIF_FORNECEDOR  varchar(100)    null,
  NOME_FORNECEDOR     varchar(300)    null,
  TP_MOV_FINANCEIRA    varchar(100)    null,
  DATA_PAG            date            null,
  NR_DL                varchar(200)    null,
  DESC_DL              varchar(200)    null,
  VL_PAGO              numeric(15,2)   null );

create table obtv_conveniente
(
  NR_MOV_FIN           bigint          null,
  NM_FAVORECIDO_OBTV_CONV varchar(300) null,
  TP_AQUISICAO         varchar(100)    null,
  VL_PAGO_OBTV_CONV    numeric(15,2)   null );

create table termo_aditivo
(
  NR_CONVENIO          bigint          null,
  NUMERO_TA            varchar(100)    null,
  TIPO_TA              varchar(100)    null,
  DT_ASSINATURA_TA     date            null,
  DT_INICIO_TA         date            null,
  DT_FIM_TA            date            null,
  JUSTIFICATIVA_TA     text            null );

create table prorroga_oficio
(
  NR_CONVENIO          bigint          null,
  NR_PRORROGA          varchar(100)    null,
  DT_INICIO_PRORROGA   date            null,
  DIAS_PRORROGA        int             null,
  DT_FIM_PRORROGA      date            null,
  DT_ASSINATURA_PRORROGA date          null,
  SIT_PRORROGA         varchar(100)    null );

create table populacao_estimada
(
  uf                   char(2)          not null,
  cd_uf                int              not null,
  cd_munic             int              not null,
  nm_munic             varchar(100)     not null,
  nr_habit             int              not null);

```

# APÊNDICE B – Comandos SQL para listagem dos dados dos convênios classificados como *outliers* e comando para geração do arquivo CSV utilizado no gráfico de *outliers*

Lista com informações adicionais dos *outliers*

select para achar os outliers

```
select x.ano, x.mes, x.nr_convênio
, x.vl_global_conv, x.sit_convênio
, x.nome_parlamentar, x.valor_repasse_emenda
, x.nm_proponente, x.desc_orgao_sup
, x.objeto_proposta, x.nome_fornecedor
, count( x.nome_fornecedor ), sum( x.vl_pago)
from (
select c.ano, c.mes, c.nr_convênio
, c.vl_global_conv, c.sit_convênio
, e.nome_parlamentar, e.valor_repasse_emenda
, p.nm_proponente , p.desc_orgao_sup
, p.objeto_proposta, pv.nome_fornecedor
, pv.vl_pago
from convenioassinado as c
join propostas_cadastradas as p
on p.id_proposta = c.id_proposta
join emendas_parlamentares as e
on c.id_proposta = e.id_proposta
left join pagamentos_favorecidos as pv
on pv.nr_convênio = c.nr_convênio
where c.nr_convênio in (
select nr_convênio
from convenioassinado as c
where c.vl_global_conv = ( 3026130.00 )
and c.vl_global_conv < (2245237.00) )
order by pv.nome_fornecedor
) as x
group by x.ano, x.mes
, x.nr_convênio, x.vl_global_conv
, x.sit_convênio, x.nome_parlamentar
, x.valor_repasse_emenda, x.nm_proponente
, x.desc_orgao_sup, x.objeto_proposta
, x.nome_fornecedor
order by x.ano, x.mes, nr_convênio, nome_parlamentar, valor_repasse_emenda, nome_fornecedor
```

### Comando para criação do arquivo com os valores dos convênios.

```
select c.ano
, c.mes
, p.uf_proponente
, c.vl_global_conv
from convenioassinado as c
join propostas_cadastradas as p
on p.id_proposta = c.id_proposta
where INSTRUMENTO_ATIVO not in ( 'NÃO', 'NAO', 'não', 'nao' )
and indassinado in ( 'SIM', 'Sim', 'sim' )
and exists ( select * from emendas_parlamentares as e where c.id_proposta = e.id_proposta and e.tipo_parlamentar = 'INDIV' )
order by c.ano, c.mes1
```

# APÊNDICE C – Comandos SQL para geração dos arquivos com os conjuntos frequentes

Comando para criação do conjunto com os dados frequentes para: Parlamentar -> fornecedor

```
select ep.nome_parlamentar
      , pv.nome_fornecedor
from public.convenioassinado as c
join public.propostas_cadastradas as pc
  on c.id_proposta = pc.id_proposta
join public.emendas_parlamentares as ep
  on ep.id_proposta = pc.id_proposta
join public.pagamentos_favorecidos as pv
  on c.nr_convenio = pv.nr_convenio
where ep.tipo_parlamentar = 'INDIVIDUAL'
      and c.ano is not null
      and c.sit_convenio is not null
      and INSTRUMENTO_ATIVO not in ( 'NÃO', 'NAO', 'não', 'nao' )
      and pc.uf_proponente = 'SE'
group by c.nr_convenio
        , ep.nome_parlamentar
        , pv.nome_fornecedor
order by c.nr_convenio
        , ep.nome_parlamentar
        , pv.nome_fornecedor
```

Comando para criação do conjunto com os dados frequentes para: Parlamentar -> Area

```
select ep.nome_parlamentar
      , pc.desc_orgao_sup
from public.convenioassinado as c
join public.propostas_cadastradas as pc
  on c.id_proposta = pc.id_proposta
join public.emendas_parlamentares as ep
  on ep.id_proposta = pc.id_proposta
where ep.tipo_parlamentar = 'INDIVIDUAL'
      and c.ano is not null
      and c.sit_convenio is not null
      and INSTRUMENTO_ATIVO not in ( 'NÃO', 'NAO', 'não', 'nao' )
      and pc.uf_proponente = 'SE'
group by c.nr_convenio
        , ep.nome_parlamentar
        , pc.desc_orgao_sup
order by 1, 2
```

Comando para criação do conjunto com os dados frequentes para: Parlamentar -> Município

```
select ep.nome_parlamentar
       , pc.nm_proponente
from public.convenioassinado as c
join public.propostas_cadastradas as pc
  on c.id_proposta = pc.id_proposta
join public.emendas_parlamentares as ep
  on ep.id_proposta = pc.id_proposta
where ep.tipo_parlamentar = 'INDIVIDUAL'
      and c.ano is not null
      and c.sit_convenio is not null
      and INSTRUMENTO_ATIVO not in ( 'NÃO', 'NAO', 'não', 'nao' )
      and pc.uf_proponente = 'SE'
group by c.nr_convenio
       , ep.nome_parlamentar
       , pc.nm_proponente
order by 1, 2
```

# APÊNDICE D – Comandos do R para estudo dos outliers

```

setwd("./files/")
conv = read.table("box_plot.csv",header=TRUE, sep = ";")
#Criando variáveis para agrupar os valores por ano.
ano2010 = conv[ which(conv$ano==2010), 6]
ano2011 = conv[ which(conv$ano==2011), 6]
ano2012 = conv[ which(conv$ano==2012), 6]
ano2013 = conv[ which(conv$ano==2013), 6]
ano2014 = conv[ which(conv$ano==2014), 6]
ano2015 = conv[ which(conv$ano==2015), 6]
ano2016 = conv[ which(conv$ano==2016), 6]
ano2017 = conv[ which(conv$ano==2017), 6]
#Gráfico com informações de convênios de todos os estados por ano e com os outliers
boxplot( ano2010, ano2011, ano2012, ano2013, ano2014, ano2015, ano2016, ano2017,
         names=c("2010", "2011", "2012", "2013", "2014", "2015", "2016", "2017") )
#Gráfico com informações de convênios de todos os estados por ano sem outliers
boxplot( ano2010, ano2011, ano2012, ano2013, ano2014, ano2015, ano2016, ano2017,
         names=c("2010", "2011", "2012", "2013", "2014", "2015", "2016", "2017"),
         outline=FALSE)
#Criando variáveis para agrupar os valores por ano e pelo estado de Sergipe.
ano2010SE = conv[ which(conv$ano==2010 & conv$uf_proponente=='SE' ), 6]
ano2011SE = conv[ which(conv$ano==2011 & conv$uf_proponente=='SE' ), 6]
ano2012SE = conv[ which(conv$ano==2012 & conv$uf_proponente=='SE' ), 6]
ano2013SE = conv[ which(conv$ano==2013 & conv$uf_proponente=='SE' ), 6]
ano2014SE = conv[ which(conv$ano==2014 & conv$uf_proponente=='SE' ), 6]
ano2015SE = conv[ which(conv$ano==2015 & conv$uf_proponente=='SE' ), 6]
ano2016SE = conv[ which(conv$ano==2016 & conv$uf_proponente=='SE' ), 6]
ano2017SE = conv[ which(conv$ano==2017 & conv$uf_proponente=='SE' ), 6]
#Gráfico com informações de convênios do estado de sergipe por ano e com os outliers
boxplot( ano2010SE, ano2011SE, ano2012SE, ano2013SE, ano2014SE, ano2015SE, ano2016SE, ano2017SE,
         names=c("2010", "2011", "2012", "2013", "2014", "2015", "2016", "2017"),
         varwidth=TRUE)
\#Gráfico com informações de convênios do estado de sergipe por ano e com os outliers
boxplot( ano2010SE, ano2011SE, ano2012SE, ano2013SE, ano2014SE, ano2015SE, ano2016SE, ano2017SE,
         names=c("2010", "2011", "2012", "2013", "2014", "2015", "2016", "2017"),
         varwidth=TRUE, outline=FALSE)
\#Obtendo a lista de os valores discrepantes por ano nos convênios no estado de Sergipe
outSE2010 = boxplot.stats(ano2013SE)$out
outSE2011 = boxplot.stats(ano2013SE)$out
outSE2012 = boxplot.stats(ano2013SE)$out
outSE2013 = boxplot.stats(ano2013SE)$out
outSE2014 = boxplot.stats(ano2013SE)$out
outSE2015 = boxplot.stats(ano2013SE)$out
outSE2016 = boxplot.stats(ano2013SE)$out
outSE2017 = boxplot.stats(ano2013SE)$out
outSE2018 = boxplot.stats(ano2013SE)$out

```



# APÊNDICE E – Comandos do R para estudo das Regras de Associação

## Preparando o ambiente

```
#Preparando o ambiente
if(! "vcd" %in% installed.packages()) install.packages("vcd", depend = TRUE)
if(! "ggplot2" %in% installed.packages()) install.packages("ggplot2" , depend = TRUE)
if(! "ggthemes" %in% installed.packages()) install.packages("ggthemes" , depend = TRUE)
if(! "gridExtra" %in% installed.packages()) install.packages("gridExtra" , depend = TRUE)
if(! "arules" %in% installed.packages()) install.packages("arules", depend = TRUE)
if(! "arulesViz" %in% installed.packages()) install.packages("arulesViz", depend = TRUE)

library(ggplot2)
library(arules)
library(arulesViz)
library(MASS)
setwd("./files/")
```

## Script para gerar o arquivo com as regras entre Parlamentar e Fornecedor

```
#importar a base para um objeto do tipo "Transaction" : Transacao_Parlamentar_Fornecedor
tPF <- read.transactions(file.choose(),format="basket",sep=";")

#Criando as regras de associacao : Regras_Parlamentar_Fornecedor
rPF <- apriori(tPF, parameter= list(supp=0.03, conf=0.6))

#Retirar as transacoes que aparecem apenas uma vez e com Lift = 1 : Regras_Limpas_Parlamentar_Fornecedor
rPF <- subset( rPF, count>1 & lift != 1 )

#retirando redundancias
rPF <- rPF[!is.redundant(rPF)]

#adicionando a medida de interesse "phi" - Coeficiente de coorelação de pearson
quality(regras) <- cbind( quality(regras), interestMeasure( regras, c("phi"), transactions = info, reuse = TRUE) )

#adicionando a medida de interesse "X-Square"
quality(regras) <- cbind( quality(regras), interestMeasure( regras, c("chiSquared"), transactions = info, reuse = TRUE) )

#adicionando a medida de interesse "p-value"
quality(regras) <- cbind( quality(regras), interestMeasure( regras, c("chiSquared"), significance = TRUE, transactions = info, reuse = TRUE) )

#Escrevendo o resultad em um arquivo
write(rPF, file = "./regras_Parlamentar_Empresa.csv", sep=";", col.names=NA)
```

## Script para gerar o arquivo com as regras entre Parlamentar e Área

```
#importar a base para um objeto do tipo "Transaction" : Transacao_Parlamentar_Area
tPA <- read.transactions(file.choose(),format="basket",sep=";")

#Criando as regras de associacao : Regras_Parlamentar_Area
rPA <- apriori(tPA, parameter= list(supp=0.002, conf=0.1))
```

```
#Retirar as transacoes que aparecem apenas uma vez e com Lift = 1 : Regras_Limpas_Parlamentar_Area
rPA <- subset( rPA, count>1 & lift != 1 )
rlPA <- subset( rPA, count>1 & lift != 1 )

#retirando redundancias
rPA <- rPA[!is.redundant(rPA)]
rlPA <- rlPA[!is.redundant(rlPA)]

\#adicionando a medida de interesse "phi" - Coeficiente de coorelação de pearson
quality(rPA) <- cbind( quality(rPA), interestMeasure( rPA, c("phi"), transactions = info, reuse = TRUE) )

\#adicionando a medida de interesse "X-Square"
quality(rPA) <- cbind( quality(rPA), interestMeasure( rPA, c("chiSquared"), transactions = info, reuse = TRUE) )

\#adicionando a medida de interesse "p-value"
quality(rPA) <- cbind( quality(rPA), interestMeasure( rPA, c("chiSquared"), significance = TRUE, transactions = info, reuse = TRUE) )

#Escrevendo o resultad em um arquivo
write(rPA, file = "./regras_Parlamentar_Area.csv", sep=";", col.names=NA)
```

### Script para gerar o arquivo com as regras entre Parlamentar e Municípios

```
#importar a base para um objeto do tipo "Transaction" : Transacao_Parlamentar_Municipio
tPM <-read.transactions(file.choose(),format="basket",sep=";")

#Criando as regras de associacao : Regras_Parlamentar_Municipio
rPM <- apriori(tPM, parameter= list(supp=0.002, conf=0.1))

#Retirar as transacoes que aparecem apenas uma vez e com Lift = 1 : Regras_Limpas_Parlamentar_Municipio
rPM <- subset( rPM, count>1 )
rPM <- subset( rPM, lift!=1 )

#retirando redundancias
rPM <- rPM[!is.redundant(rPM)]

\#adicionando a medida de interesse "phi" - Coeficiente de coorelação de pearson
quality(rPM) <- cbind( quality(rPM), interestMeasure( rPM, c("phi"), transactions = info, reuse = TRUE) )

\#adicionando a medida de interesse "X-Square"
quality(rPM) <- cbind( quality(rPM), interestMeasure( rPM, c("chiSquared"), transactions = info, reuse = TRUE) )

\#adicionando a medida de interesse "p-value"
quality(rPM) <- cbind( quality(rPM), interestMeasure( rPM, c("chiSquared"), significance = TRUE, transactions = info, reuse = TRUE) )

#Escrevendo o resultad em um arquivo
write(rPM, file = "./regras_Parlamentar_Municipio.csv", sep=";", col.names=NA)
```

# APÊNDICE F – Scripts em Python

Script para efetuar o download dos arquivos de dados utilizados.

```
import urllib
import urllib2
import time
import os

from datetime import datetime

listaUrl = [ 'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_programa.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_programa_proposta.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_proposta.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_convenio.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_emenda.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_plano_aplicacao.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_empenho.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_desembolso.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_pagamento.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_obtv_conveniente.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_historico_situacao.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_ingresso_contrapartida.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_termo_aditivo.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_prorroga_oficio.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_meta_crono_fisico.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_etapa_crono_fisico.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_consortorios.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/csv/siconv_empenho_desembolso.csv.zip',
'http://portal.convenios.gov.br/images/docs/CGSIS/Modelo_de_Dados_SICONV.jpg',
'http://portal.convenios.gov.br/images/docs/CGSIS/Historico_de_versoes.pdf',
'http://portal.convenios.gov.br/images/docs/CGSIS/siconv_dicionario_dados.pdf']

print "baixando com urllib"
path = '/Users/othon.campos/Documents/siconv'
now = datetime.now()
pasta = str( now.year ).zfill(4) + '.' + str( now.month ).zfill(2) + '.' + str( now.day ).zfill(2)
path += "/" + pasta

if ( not os.path.isdir( path ) ):
os.mkdir(path)

fileName = path + '/' + os.path.basename( listaUrl[0] )
tempo0 = time.time()
i = 0
while( i < len(listaUrl) ):
url = listaUrl[i]
fileName = path + '/' + os.path.basename( url )
tempo10 = time.time()
urllib.urlretrieve(url, fileName)
tempo11 = time.time()
print "Tempo: " + fileName + ' : ' + str( tempo11 - tempo10)
i+=1
tempo1 = time.time()
print "Tempo final: " + str( tempo1 - tempo0)
```

## Script para ajustar os ";" presentes no texto.

```
import csv
import string
import sys

def imprimirLog(texto, permite):
    if permite:
        print(texto)

def to_bool(value):
    valid = {'true': True, 't': True, '1': True,
            'false': False, 'f': False, '0': False,}

    if isinstance(value, bool):
        return value

    if not isinstance(value, basestring):
        raise ValueError('invalid literal for boolean. Not a string.')

    lower_value = value.lower()
    if lower_value in valid:
        return valid[lower_value]
    else:
        raise ValueError('invalid literal for boolean: "%s" % value)

nome_arquivo = sys.argv[1]
imprimir_log = to_bool(sys.argv[2])
path = path = './'
fileName = path + '/' + nome_arquivo
arq = open(fileName, 'r')
texto = arq.readlines()
newFileName = fileName+".csv"
arqFim = open(newFileName, 'w')

imprimirLog( "Verificando o arquivo: " + nome_arquivo, imprimir_log)
cont = 0

for linha in texto:
    fim = 0
    verificarAspas = True
    if linha.find('null') > 0:
        linha = linha.replace('null', '')

    while( verificarAspas ):
        if fim == 0 :
            inicio = linha.find(';')
        elif fim < len(linha) :
            aux = linha[fim+2:]
            inicio = aux.find(';')
            if inicio > 0:
                inicio = inicio + fim + 2
            else:
                verificarAspas = False
                break

        if inicio > 0:
            aux = linha[inicio+2:]
            fim = aux.find(';') + inicio + 2
```

```

imprimirLog(str(inicio) + ":" + str(fim), imprimir_log)
aux0 = linha[inicio+1:fim]
imprimirLog(aux0, imprimir_log)
aux1 = aux0.replace(';', '|')
imprimirLog(aux1, imprimir_log)
imprimirLog(linha, imprimir_log)
linha = linha.replace(aux0, aux1)
imprimirLog(linha, imprimir_log)
cont+=1
else:
verificarAspas = False

arqFim.write(linha)
imprimirLog("Novo Arquivo criado: " + newFileName, imprimir_log)
arq.close()
arqFim.close()

```

Script para ler os arquivos CSVs e inserir os dados no banco de dados.

```

#!/usr/bin/env python3
# -*- coding: utf-8 -*-

import sys
import mysql.connector
import csv
import locale
from unicodedata import normalize

def main(user, db, table, csvfile):
    try:
        conn = getconn(user, db,)
    except MySQLdb.Error as e:
        print ("Error %d: %s" % (e.args[0], e.args[1]) )
        sys.exit (1)
    loadcsv(conn, table, csvfile)
    conn.close()

def remover_acentos(txt, codif='utf-8'):
    return normalize('NFKD', txt.decode(codif)).encode('ASCII', 'ignore')

def getconn(user, db, passwd=""):
    config = {
        'user': 'root',
        'password': 'P@ssw0rd',
        'host': '127.0.0.1',
        'database': 'tt_viagens',
        'raise_on_warnings': True}
    conn = mysql.connector.connect(**config)
    return conn

def nullify(L):
    def f(x):
        if(x == ""):
            return None
        else:
            return x
    return [f(x) for x in L]

def efetuarCommit(conn, qtd_insert, qtd_linha)

```

```

print('Committing '%s' rows to database ... processed '%s' rows of file ' % ( qtd_insert, qtd_linha ))
conn.commit()

def loadcsv(conn, table, filename):
    nameFileErro = filename + '.err'
    nameFileRej = filename + '.rej'
    fileErro = open( nameFileErro, 'w')
    fileRej = open( nameFileRej , 'w')

    cursor = conn.cursor()
    with open(filename, newline='', encoding='latin-1') as file:
        f = csv.reader(file, delimiter=';')
        header = f.__next__()
        numfields = len(header)
        dadosQuery = buildInsertCmd(cursor, table, numfields)
        query = dadosQuery["query"]
        listTypeQuery = dadosQuery["listType"]
        litsDecimal=[]
        i=0
        for j in listTypeQuery:
            if( str(j).find('decimal') > 0 ):
                litsDecimal.append( i )
            i+=1
        max_inserts = 10000
        qtd_linha = 1
        qtd_insert=0

        for line in f:
            try:
                vals = nullify(line)
                vals.insert(0, "2018")
                for j in litsDecimal:
                    valor = vals[j].replace(',','.')
                    valor = '{:0>15}'.format(valor)
                    vals[j] = valor
                cursor.execute(query, vals)
                qtd_insert+=1
                if qtd_insert % max_inserts == 0:
                    print('Committing '%s' rows to database ... processed '%s' rows of file ' % ( qtd_insert, qtd_linha ))
                    conn.commit()
            except Exception as e:
                fileErro.write(str(e))
                fileErro.write("%s : %s\n" %(str(qtd_insert), line))
                fileRej.write("%s\n" %line)
            print('Committing '%s' rows to database ... processed '%s' rows of file ' % ( qtd_insert, qtd_linha ))
            conn.commit()
            qtd_linha+=1
            print('Committing '%s' rows to database ... processed '%s' rows of file ' % ( qtd_insert, qtd_linha ))
            conn.commit()
            cursor.close()
            fileErro.close()
            fileRej.close()
        return

def buildInsertCmd(cursor, table, numfields):
    query = 'show fields from %s' %table
    cursor.execute( query )
    list = cursor.fetchall()
    listColumns =[]
    listTypes =[]

```

```
for l in list:
    if( l[5] != 'auto_increment'):
        listColumns.append(l[0])
        listTypes.append(l[1])

numfields = len(listColumns)
assert(numfields > 0)
placeholders = (numfields-1) * "%s, " + "%s"

arguments = ', '.join(listColumns)
query = ("insert into %s" % table) + ("( %s )" % arguments) + (" values (%s)" % placeholders)
dadosQuery = {"query": query, "listType": listTypes }
return dadosQuery

if __name__ == '__main__':
    args = sys.argv[1:]
    if(len(args) < 4):
        print ("error: arguments: user db table csvfile")
        sys.exit(1)
    main(*args)
```

# **Anexos**



## **ANEXO A – Dicionário de dados**

Programas Disponibilizados	
NOME DO CAMPO	DESCRIÇÃO
COD_ORGAO_SUP_PROGRAMA	Código do Órgão executor do Programa
DESC_ORGAO_SUP_PROGRAMA	Nome do Órgão executor do Programa
ID_PROGRAMA	Código Sequencial do Sistema para um Programa
COD_PROGRAMA	Chave que identifica o programa composta por: (Cód.Órgão+Ano+Cód.Sequencial do Sistema)
NOME_PROGRAMA	Descrição do Programa de Governo
SIT_PROGRAMA	Situação atual do Programa. Domínio: Cadastrado; Disponibilizado; Inativo
DATA_DISPONIBILIZACAO	Data de disponibilização do Programa
ANO_DISPONIBILIZACAO	Ano de disponibilização do Programa
DT_PROG_INI_RECEB_PROP	Data Início para o recebimento das propostas voluntárias para o Programa
DT_PROG_FIM_RECEB_PROP	Data Fim para o recebimento das propostas voluntárias para o Programa
DT_PROG_INI_EMENDA_PAR	Data Início para o recebimento das propostas de Emenda Parlamentar para o Programa
DT_PROG_FIM_EMENDA_PAR	Data Fim para o recebimento das propostas de Emenda Parlamentar para o Programa
DT_PROG_INI_BENEF_ESP	Data Início para o recebimento das propostas de beneficiário específico para o Programa
DT_PROG_FIM_BENEF_ESP	Data Fim para o recebimento das propostas de beneficiário específico para o Programa
MODALIDADE_PROGRAMA	Modalidade do Programa. Domínio pode ser: CONTRATO DE REPASSE, CONVENIO, TERMO DE COLABORACAO, TERMO DE FOMENTO e TERMO DE PARCERIA
NATUREZA_JURIDICA_PROGRAMA	Natureza Jurídica Atendida pelo Programa. Domínio: Administração Pública Estadual ou do Distrito Federal, Administração Pública Municipal, Consórcio Público, Empresa pública/Sociedade de economia mista e Organização da Sociedade Civil
UF_PROGRAMA	Ufs Habilitadas para o Programa. Domínio: AC, AL, AM, AP, BA, CE, DF, ES, GO, MA, MG, MS, MT, PA, PB, PE, PI, PR, RJ, RN, RO, RR, RS, SC, SE, SP, TO, <null>. Quando o valor é nulo, o programa atende a todo o Brasil
ACAO_ORCAMENTARIA	Número da Ação Orçamentária

Programas/Propostas	
NOME DO CAMPO	DESCRIÇÃO
ID_PROGRAMA	Código Sequencial do Sistema para um Programa
ID_PROPOSTA	Código Sequencial do Sistema para uma Proposta

Propostas Cadastradas	
NOME DO CAMPO	DESCRIÇÃO
ID_PROPOSTA	Código Sequencial do Sistema para uma Proposta
UF_PROponente	UF do Proponente. Domínio: AC, AL, AM, AP, BA, CE, DF, ES, GO, MA, MG, MS, MT, PA, PB, PE, PI, PR, RJ, RN, RO, RR, RS, SC, SE, SP, TO
MUNIC_PROponente	Município do Proponente
COD_ORGAO_SUP	Código do Órgão Superior do Concedente
DESC_ORGAO_SUP	Nome do Órgão Superior do Concedente
NATUREZA_JURIDICA	Natureza Jurídica do Proponente. Domínio: Administração Pública Estadual ou do Distrito Federal, Administração Pública Municipal, Consórcio Público, Empresa pública/Sociedade de economia mista e Organização da Sociedade Civil
NR_PROPOSTA	Número da Proposta gerado pelo Siconv
DIA_PROP	Dia do cadastro da Proposta
MES_PROP	Mês do cadastro da Proposta
ANO_PROP	Ano do cadastro da Proposta
DIA_PROPOSTA	Data do cadastro da Proposta
COD_ORGAO	Código do Órgão ou Entidade Concedente
DESC_ORGAO	Nome do Órgão ou Entidade Concedente
MODALIDADE	Modalidade da Proposta. Domínio pode ser: CONTRATO DE REPASSE, CONVENIO, TERMO DE COLABORACAO, TERMO DE FOMENTO e TERMO DE PARCERIA
IDENTIF_PROponente	CNPJ do Proponente
NM_PROponente	Nome da Entidade Proponente
CEP_PROponente	CEP do Proponente
ENDERECO_PROponente	Endereço do Proponente
BAIRRO_PROponente	Bairro do Proponente
NM_BANCO	Nome do Banco para depósito do recurso da Transferência Voluntária
SITUACAO_CONTA	Situação atual da conta bancária do instrumento. Domínio: Aguardando Retorno do Banco, Enviada, Cadastrada, Registrada, Erro na Abertura de Conta, Regularizada, A Verificar, Aguardando Envio e Pendente de Regularização
SITUACAO_PROJETO_BASICO	Situação atual do Projeto Básico/Termo de Referência. Domínio: Aguardando Projeto Básico, Não Cadastrado, Projeto Básico Aprovado, Projeto Básico em Análise, Projeto Básico em Complementação, Projeto Básico Rejeitado
SIT_PROPOSTA	Situação atual da Proposta. Domínio pode ser: Proposta/Plano de Trabalho Cadastrados, Proposta/Plano de Trabalho em Análise, Proposta/Plano de Trabalho Rejeitados, Proposta/Plano de Trabalho Aprovados, etc
DIA_INIC_VIGENCIA_PROPOSTA	Data Início da Vigência da Proposta
DIA_FIM_VIGENCIA_PROPOSTA	Data Fim da Vigência da Proposta
OBJETO_PROPOSTA	Descrição do Objeto da Proposta
VL_GLOBAL_PROP	Valor Global da proposta cadastrada (Valor de Repasse Proposta + Valor Contrapartida Proposta)
VL_REPASSE_PROP	Valor de Repasse do Governo Federal referente a proposta cadastrada
VL_CONTRAPARTIDA_PROP	Valor da Contrapartida apresentada na proposta pelo conveniente

Emendas Parlamentares	
NOME DO CAMPO	DESCRIÇÃO
ID_PROPOSTA	Código Sequencial do Sistema para uma Proposta
QUALIF_PROponente	Qualificação do proponente
COD_PROGRAMA_EMENDA	Chave que identifica o programa composta por: (Cód.Órgão+Ano+Cód.Sequencial do Sistema)
NR_EMENDA	Número da Emenda Parlamentar
NOME_PARLAMENTAR	Nome do Parlamentar
BENEFICIARIO_EMENDA	CNPJ do Proponente
IND_IMPOSITIVO	Indicativo de Orçamento Impositivo (Tipo Parlamentar igual a INDIVIDUAL + Ano de Cadastro da Proposta >= 2014). Domínio: SIM, NÃO
TIPO_PARLAMENTAR	Tipo do Parlamentar. Domínio pode ser: INDIVIDUAL, COMISSAO, BANCADA
VALOR_REPASSE_PROPOSTA_EMENDA	Valor da Emenda cadastrada na proposta
VALOR_REPASSE_EMENDA	Valor da Emenda assinada

Histórico Situação Propostas/Convênios	
NOME DO CAMPO	DESCRIÇÃO
ID_PROPOSTA	Código Sequencial do Sistema para uma Proposta
NR_CONVENIO	Número gerado pelo Siconv. Possui faixa de numeração reservada que vai de 700000 a 999999
DIA_HISTORICO_SIT	Data de entrada da situação no sistema
HISTORICO_SIT	Situação histórica da Proposta/Convênio
DIAS_HISTORICO_SIT	Dias em que a Proposta/Convênio permaneceu na situação
COD_HISTORICO_SIT	Código da situação histórica da Proposta/Convênio, contendo a ordem cronológica do ciclo de vida de um convênio

Plano de Aplicação Detalhado	
NOME DO CAMPO	DESCRIÇÃO
ID_PROPOSTA	Código Sequencial do Sistema para uma Proposta
SIGLA	UF cadastrada referente a localidade do item. Domínio: AC, AL, AM, AP, BA, CE, DF, ES, GO, MA, MG, MS, MT, PA, PB, PE, PI, PR, RJ, RN, RO, RR, RS, SC, SE, SP, TO
MUNICIPIO	Município cadastrado referente a localidade do item
NATUREZA_AQUISICAO	Código de natureza de aquisição
DESCRICAO_ITEM	Descrição do Item
CEP_ITEM	CEP cadastrado referente a localidade do item
ENDERECO_ITEM	Endereço cadastrado referente a localidade do item
TIPO_DESPESA_ITEM	Tipo da Despesa. Domínio: SERVICO, BEM, OUTROS, TRIBUTOS, OBRA e DESPESA ADMINISTRATIVA
NATUREZA_DESPESA	Natureza da Despesa referente ao item
SIT_ITEM	Situação atual do Item. Domínio: APROVADO
QTD_ITEM	Quantidade de Itens
VALOR_UNITARIO_ITEM	Valor unitário do item
VALOR_TOTAL_ITEM	Valor total do item

Convênios Assinados	
NOME DO CAMPO	DESCRIÇÃO
NR_CONVENIO	Número gerado pelo Siconv. Possui faixa de numeração reservada que vai de 700000 a 999999
ID_PROPOSTA	Código Sequencial do Sistema para uma Proposta
DIA	Data em que o Convênio foi assinado
MES	Mês em que o Convênio foi assinado
ANO	Ano Assinatura do Convênio
DIA_ASSIN_CONV	Data Assinatura do Convênio
SIT_CONVENIO	Situação atual do Convênio. Domínio: Em execução, Convênio Anulado, Prestação de Contas enviada para Análise, Prestação de Contas Aprovada, Prestação de Contas em Análise, Prestação de Contas em Complementação, Inadimplente, etc
SUBSTITUACAO_CONV	Sub-Situação atual do Convênio. Domínio: Convênio, Convênio Cancelado, Convênio Encerrado, Proposta, Em aditativação
SITUACAO_PUBLICACAO	Situação atual da Publicação do instrumento. Domínio: Publicado e Transferido para IN
INSTRUMENTO_ATIVO	Convênios que ainda não foram finalizados. Domínio: SIM, NÃO
IND_OPERA_OBTV	Indicativo de que o Convênio opera com OBTV. Domínio: SIM, NÃO
NR_PROCESSO	Número interno do processo físico do instrumento
UG_EMITENTE	Número da Unidade Gestora
DIA_PUBL_CONV	Data da Publicação do Convênio
DIA_INIC_VIGENC_CONV	Data de Início de Vigência do Convênio
DIA_FIM_VIGENC_CONV	Data de Fim de Vigência do Convênio
DIAS_PREST_CONTAS	Pazo para a Prestação de Contas do Convênio
DIA_LIMITE_PREST_CONTAS	Data limite para Prestação de Contas do Convênio
QTDE_CONVENIOS	Quantidade de Instrumentos Assinados
QTD_TA	Quantidade de Termos Aditivos
QTD_PRORROGA	Quantidade de Prorrogas de Ofício
VL_GLOBAL_CONV	Valor global dos Instrumentos assinados (Valor de Repasse + Valor Contrapartida)
VL_REPASSE_CONV	Valor total do aporte do Governo Federal referente a celebração do Instrumento
VL_CONTRAPARTIDA_CONV	Valor total da Contrapartida que será disponibilizada pelo conveniente
VL_EMPENHADO_CONV	Valor total empenhado do Governo Federal para os Instrumentos
VL_DESEMBOLSADO_CONV	Valor total desembolsado do Governo Federal para a conta do Instrumento
VL_SALDO_REMAN_TESOURO	Valores devolvidos ao Tesouro ao término do instrumento
VL_SALDO_REMAN_CONVENIENTE	Valores devolvidos ao Conveniente ao término do instrumento
VL_RENDIMENTO_APLICACAO	Valores referentes aos rendimentos de aplicação financeira
VL_INGRESSO_CONTRAPARTIDA	Total de valores referente a ingresso de contrapartida dos Instrumentos

Empenhos Realizados	
NOME DO CAMPO	DESCRIÇÃO
NR_CONVENIO	Número gerado pelo Siconv. Possui faixa de numeração reservada que vai de 700000 a 999999
NR_EMPENHO	Número da Nota de Empenho
TIPO_NOTA	Código do Tipo de Empenho
DESC_TIPO_NOTA	Descrição do Tipo de Empenho. Domínio: Empenho Original, Empenho de Despesa Pré-Empenhada, Anulação de Empenho, Reforço de Empenho, Estorno de Anulação de Empenho, etc
DATA_EMISSAO	Data de emissão do Empenho
COD_SITUACAO_EMPENHO	Código da Situação atual do empenho
DESC_SITUACAO_EMPENHO	Descrição da Situação atual do empenho. Domínio: Registrado no SIAFI e Enviado
VALOR_EMPENHO	Valor empenhado

Desembolsos (Ordem Bancária)	
NOME DO CAMPO	DESCRIÇÃO
NR_CONVENIO	Número gerado pelo Siconv. Possui faixa de numeração reservada que vai de 700000 a 999999
DT_ULT_DESEMBOLSO	Data da última Ordem Bancária gerada
QTD_DIAS_SEM_DESEMBOLSO	Indicador de dias sem desembolso. Domínio: 90,180 e 365 dias
DATA_DESEMBOLSO	Data da Ordem Bancária
ANO_DESEMBOLSO	Ano da Ordem Bancária
MES_DESEMBOLSO	Mês da Ordem Bancária
NR_SIAFI	Número do Documento no SIAFI
VL_DESEMBOLSADO	Valor disponibilizado pelo Governo Federal para a conta do instrumento

Ingresso de Contrapartida	
NOME DO CAMPO	DESCRIÇÃO
NR_CONVENIO	Número gerado pelo Siconv. Possui faixa de numeração reservada que vai de 700000 a 999999
DT_INGRESSO_CONTRAPARTIDA	Data da disponibilização do recurso por parte do Conveniente
VL_INGRESSO_CONTRAPARTIDA	Valor disponibilizado pelo Conveniente para a conta do instrumento

Pagamentos a Favorecidos	
NOME DO CAMPO	DESCRIÇÃO
NR_MOV_FIN	Número identificador da movimentação financeira
NR_CONVENIO	Número gerado pelo Siconv. Possui faixa de numeração reservada que vai de 700000 a 999999
IDENTIF_FORNECEDOR	CNPJ (CPF do Fornecedor)
NOME_FORNECEDOR	Nome do Fornecedor
TP_MOV_FINANCEIRA	Tipo da movimentação financeira realizada. Domínio: Pagamento a favorecido, Pagamento a favorecido com OBTV
DATA_PAG	Data da realização do pagamento
NR_DL	Número identificador do Documento de Liquidação
DESC_DL	Descrição do Documento de Liquidação. Domínio: DIÁRIAS, DUPLICATA, FATURA, FOLHA DE PAGAMENTO, NOTA FISCAL, NOTA FISCAL / FATURA, OBTV PARA EXECUTOR, OBTV PARA O CONVENIENTE, etc
VL_PAGO	Valor do pagamento

OBTV para Conveniente	
NOME DO CAMPO	DESCRIÇÃO
NR_MOV_FIN	Número identificador da movimentação financeira
IDENTIF_FAVORECIDO_OBTV_CONV	CNPJ/CPF do Favorecido recebedor do pagamento
NM_FAVORECIDO_OBTV_CONV	Nome do Favorecido recebedor do pagamento
TP_AQUISICAO	Tipo de Aquisição
VL_PAGO_OBTV_CONV	Valor pago ao favorecido

Termo Aditivo	
NOME DO CAMPO	DESCRIÇÃO
NR_CONVENIO	Número gerado pelo Siconv. Possui faixa de numeração reservada que vai de 700000 a 999999
NUMERO_TA	Número do Termo Aditivo
TIPO_TA	Tipo do Termo Aditivo
VL_GLOBAL_TA	Valor Global referente ao TA
VL_REPASSE_TA	Valor de Repasse referente ao TA
VL_CONTRAPARTIDA_TA	Valor de Contrapartida referente ao TA
DT_ASSINATURA_TA	Data da assinatura do Termo Aditivo
DT_INICIO_TA	Data Início de Vigência do Termo Aditivo
DT_FIM_TA	Data Fim de Vigência do Termo Aditivo
JUSTIFICATIVA_TA	Justificativa para a realização do Termo Aditivo

Prorroga de Ofício	
NOME DO CAMPO	DESCRIÇÃO
NR_CONVENIO	Número gerado pelo Siconv. Possui faixa de numeração reservada que vai de 700000 a 999999
NR_PRORROGA	Número do Prorroga de Ofício
DT_INICIO_PRORROGA	Data Início de Vigência do Prorroga de Ofício
DIAS_PRORROGA	Dias de prorrogação
DT_FIM_PRORROGA	Data Fim de Vigência do Prorroga de Ofício
DT_ASSINATURA_PRORROGA	Data de assinatura do Prorroga de Ofício
SIT_PRORROGA	Situação atual do Prorroga de Ofício. Domínio: DISPONIBILIZADA, PUBLICADA

## **ANEXO B – Lista dos *outliers***

Ano	Nr. Conv.	Vl. Global	Proponente	Área
2013	784078	R\$2.342.275,85	Mun. de Capela	Min. do Esporte
Descrição		Implantação e modernização de infraestrutura esportiva.		
Parlamentar	Nome		Valor Emenda	
	Laercio Oliveira		R\$ 975.000,00	
	Eduardo Amorim		R\$ 682.500,00	
Favorecido	Nome		Qtd.	Vl. Fatura
	DICON - Engenharia LTDA - EPP		18	R\$1.580.439,23
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2013	784264	R\$ 2.245.235,86	Mun. de Rosário do Catete	Min. do Esporte
Descrição		Implantação e modernização de infraestrutura esportiva.		
Parlamentar	Nome		Valor Emenda	
	Rogerio Carvalho		R\$ 1.560.000,00	
Favorecido	Nome		Qtd.	Vl. Fatura
	Mouro Engenharia LTDA - EPP		1	R\$ 82.040,17
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2013	798585	R\$ 1.741.250,00	Mun. de Capela	Min. do Esporte
Descrição		Implantação e Modernização de Infraestrutura para o Esporte Educacional e de Lazer do Município de Capela/SE.		
Parlamentar	Nome		Valor Emenda	
	Andre Moura		R\$ 1.706.250,00	
Favorecido	Nome		Qtd.	Vl. Fatura
	Vespe Construções LTDA - EPP		1	R\$ 55.871,45
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2014	809215	R\$ 3.026.130,00	Mun. de Frei Paulo	Min. das Cidades
Descrição		Implantação de recapeamento e pav. em vias públicas urbanas de Frei Paulo.		
Parlamentar	Nome		Valor Emenda	
	Almeida Lima		R\$ 2.965.600,00	
Favorecido	Nome		Qtd.	Vl. Fatura
	Torre Empreend. Rural e Construção LTDA		2	R\$1.950.404,19
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2015	814910	R\$2.943.000,00	Assoc.Hospitalar de Sergipe	Min. da Saúde
Descrição		Aquisição de equipamento e material permanente para unidade de atenção especializada em saúde		
Parlamentar	Nome		Valor Emenda	
	Fabio Reis		R\$2.943.000,00	
Favorecido	Nome		Qtd.	Vl. Fatura
	ALFA MED SISTEMAS MEDICOS LTDA		1	R\$90.000,00
	ALL MED PROD. HOSP. LTDA – ME		3	R\$105.000,00
	BIODIAG DIAG. E HOSP. LTDA – EPP		1	R\$5.000,00
	CARDIO SISTEMAS COMERCIAL E INDUSTRIAL LTDA		1	R\$51.400,00
	HOSPLIFE COMERCIO DE EQUIP. HOSPITALARES LTDA		1	R\$84.000,00
	KSS COMERCIO E INDUSTRIA DE EQUIPAMENTOS MEDICO LTDA		1	R\$170.550,00
	LABOR MED APARELHAGEM DE PRECISAO LTDA		1	R\$400.000,00
	LIFE COMERCIO E SERVICOS EIRELI		5	R\$706.600,00
	LUSMED COMERCIO DE PRODUTOS MEDICOS HOSPITALARES LTDA		2	R\$7.700,00
	MAGNAMED TECNOLOGIA MEDICA S/A		1	R\$65.000,00
	MAXIMA COMERCIO, REPRESENTACOES E ASSISTENCIA TECNICA LTDA - EPP		1	R\$10.000,00
	MEDICALWAY EQUIP. MEDICOS LTDA		1	R\$146.000,00
	PHILIPS MEDICAL SYSTEMS LTDA		1	R\$70.000,00

Ano	Nr. Conv.	Vl. Global	Proponente	Área
2015	821488	R\$ 1.993.722,72	MUN. DE ARACAJU	MIN. DAS CIDADES
Descrição		IMPLANTAÇÃO DE PAVIMENTAÇÃO NA AVENIDA A - BAIRRO JAPAOZINHO - ARACAJU-SE.		
Parlamentar		Nome	Valor Emenda	
		MARIA DO CARMO ALVES	R\$ 1.949.022,72	
Favorecido		Nome	Qtd.	Vl. Fatura
		AMT PROJETOS E SERVICOS LTDA	3	R\$ 87.390,78
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2015	819859	R\$ 2.036.000,00	MUN. DE ARACAJU	MIN. DAS CIDADES
Descrição		IMPLANTAÇÃO DE PAVIMENTAÇÃO DE RUAS NO LOTEAMENTO MOEMA MEIRE - 1 ETAPA, BAIRRO SANTOS DUMONT, SITUADO NO MUNICÍPIO DE ARACAJU/SE.		
Parlamentar		Nome	Valor Emenda	
		MENDONCA PRADO	R\$ 1.976.600,00	
Favorecido		Nome	Qtd.	Vl. Fatura
		AMT PROJETOS E SERVICOS LTDA	10	R\$ 408.506,42
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2015	825740	R\$ 2.222.300,00	ESTADO DE SERGIPE	MIN. INT. NACIONAL
Descrição		Fortalecimento da Agricultura Familiar em projetos de assentamentos da reforma agrária, comunidades tradicionais e no Estado de Sergipe.		
Parlamentar		Nome	Valor Emenda	
		JOAO DANIEL	R\$ 2.200.000,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2015	822376	R\$ 3.912.387,00	ESTADO DE SERGIPE	MIN. INT. NACIONAL
Descrição		Construção de uma ponte sobre o Rio Sergipe, Trecho: Entrada SE-230/Povoado Taboca, no Mun. de Nossa Senhora das Dores, no Estado de Sergipe, visando facilitar o fluxo de escoamento dos produtos resultantes dos arranjos produtivos locais.		
Parlamentar		Nome	Valor Emenda	
		ALMEIDA LIMA	R\$ 3.908.474,61	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2015	824897	R\$ 3.952.638,90	FUNDO MUN. SAUDE	MIN. DA SAUDE
Descrição		AMPLIAÇÃO DE UNIDADE DE ATENÇÃO ESPECIALIZADA EM SAÚDE.		
Parlamentar		Nome	Valor Emenda	
		MARIA DO CARMO ALVES	R\$ 3.952.638,90	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2016	829748	R\$ 2.331.252,34	MUN. DE ITABAIANA	MIN. DAS CIDADES
Descrição		Pavimentação de Ruas no Bairro Dr. José Milton Machado e outros		
Parlamentar		Nome	Valor Emenda	
		EDUARDO AMORIM	R\$ 1.976.600,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2016	823854	R\$ 1.251.500,00	ESTADO DE SERGIPE	MIN. DA CULTURA
Descrição		Rec. das fachadas e telhado da Catedral Metropolitana de Aracaju, incluindo as Instalações Eletrônicas sobre o forro e SPDA.		
Parlamentar		Nome	Valor Emenda	
		ANTONIO CARLOS VALADARES	R\$ 250.000,00	
		EDUARDO AMORIM	R\$ 500.000,00	
		LAERCIO OLIVEIRA	R\$ 250.000,00	
		MARIA DO CARMO ALVES	R\$ 250.000,00	
Favorecido		Nome	Qtd.	Vl. Fatura
		ESSENCIAL TRANSP. & CONSTR. LTDA	12	R\$ 1.029.025,29

Ano	Nr. Conv.	Vl. Global	Proponente	Área
2016	833214	R\$ 1.868.459,29	ESTADO DE SERGIPE	SEC.ESPEC.DE AGRIC.FAMIL.E DO DESENV.AGRARIO
Descrição		Aquisição de equipamentos para fortalecimento da agricultura familiar, a saber: balança rodoviária e respectiva obra de instalação  trator agrícola  carreta agrícola  grade aradora  Kit de irrigação instalado  Câmara Frigorífica.		
Parlamentar		Nome	Valor Emenda	
		JOAO DANIEL	R\$ 1.849.774,70	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2016	828874	R\$ 1.569.047,02	MUN. DE ARACAJU	MIN. DAS CIDADES
Descrição		IMPLANTAÇÃO DE PAVIMENTAÇÃO E DRENAGEM NA ZONA DE EXPANSÃO NAS AV. Coronel Renir Reis Damasceno., AV. Dr. José Maria Rodrigues santos, Rua Alameda dos Abaí , Rua José Carlos Vital, Rua Augusto da Silva, Rua Gilmar Moraes Amaral, Rua SGF, Rua H, Rua G, Rua N, Rua Antonio Prado Barros E Rua Maria de Lurdes NO MUNICÍPIO DE ARACAJU.		
Parlamentar		Nome	Valor Emenda	
		ADELSON BARRETO	R\$1.530.777,59	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	834377	R\$ 1.695.416,00	ESTADO DE SERGIPE	MIN. DA CULTURA
Descrição		Segunda Etapa da Restauração da Catedral Metropolitana do Estado de Sergipe, em Aracaju/SE.		
Parlamentar		Nome	Valor Emenda	
		ANTONIO CARLOS VALADARES	R\$ 250.000,00	
		EDUARDO AMORIM	R\$ 443.716,00	
		LAERCIO OLIVEIRA	R\$ 250.000,00	
		MARIA DO CARMO ALVES	R\$ 500.000,00	
		VALADARES FILHO	R\$ 250.000,00	
Favorecido		Nome	Qtd.	Vl. Fatura
		MGB EXECUCOES EIRELI - ME	5	R\$ 544.661,89
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	846039	R\$ 1.183.543,08	ESTADO DE SERGIPE	MIN. DAS CIDADES
Descrição		Obras e Serviços de pavimentação granítica em ruas dos Povoados Rio Fundo, Mariquita e Pista do Pau Grande, no município de Lagarto, no Estado de Sergipe.		
Parlamentar		Nome	Valor Emenda	
		FABIO REIS	R\$ 591.771,54	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	844059	R\$ 500.000,00	MUN. DE N S DA GLORIA	MIN. DO DESENV. SOCIAL
Descrição		Construção de Centro de CRAS		
Parlamentar		Nome	Valor Emenda	
		VIRGINIO DE CARVALHO	R\$ 500.000,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	845269	R\$ 500.000,00	MUN. DE SIRIRI	MIN. DAS CIDADES
Descrição		PAVIMENTAÇÃO DE VIAS PUBLICAS DO MUN. DE SIRIRI.		
Parlamentar		Nome	Valor Emenda	
		JOAO BOSCO COSTA	R\$ 493.100,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	846311	R\$ 592.600,00	MUN. DE LARANJEIRAS	MIN. DAS CIDADES
Descrição		Pavimentação e Drenagem de ruas no Município de Laranjeiras\SE		
Parlamentar		Nome	Valor Emenda	
		ADELSON BARRETO	R\$ 592.000,00	



Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	844978	R\$ 504.000,00	MUN. DE ESTANCIA	MIN. DAS CIDADES
Descrição		Perfuração de 10 (dez) Poços Artesianos no Município de Estância.		
Parlamentar		Nome	Valor Emenda	
		ANDRE MOURA	R\$ 502.760,55	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	845544	R\$ 500.000,00	MUN. DE BOQUIM	MIN. DAS CIDADES
Descrição		Pavimentação de diversas ruas no município de Boquim.		
Parlamentar		Nome	Valor Emenda	
		LAERCIO OLIVEIRA	R\$493.100,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	844927	R\$ 800.000,00	MUN. ITAPORANGA D'AJUDA	MIN. DAS CIDADES
Descrição		Implantação/melhoria de infraestrutura para mobilidade urbana no município de Itaporanga D'Ajuda.		
Parlamentar		Nome	Valor Emenda	
		VIRGINIO DE CARVALHO	R\$ 789.800,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	852158	R\$ 500.000,00	MUN. N S DA GLORIA	MIN. DO ESPORTE
Descrição		Construção de uma quadra coberta no povoado Quixaba		
Parlamentar		Nome	Valor Emenda	
		JONY MARCOS	R\$ 487.500,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	852190	R\$ 488.500,00	MUN. DE ESTANCIA	MIN. DO ESPORTE
Descrição		Construção de Quadra Poliesportiva no Município de Estância		
Parlamentar		Nome	Valor Emenda	
		JONY MARCOS	R\$ 487.500,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	843981	R\$ 1.000.000,00	ESTADO DE SERGIPE	MIN. DO ESPORTE
Descrição		Reforma e ampliação da infraestrutura esportiva do Estado de Sergipe		
Parlamentar		Nome	Valor Emenda	
		Fabio Mitidieri	R\$ 975.000,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	844019	R\$ 629.147,05	Mun. de Aquidabã	Min. do Esporte
Descrição		Implantação e Modernização de Infraestrutura Esportiva no Município de Aquidabã.		
Parlamentar		Nome	Valor Emenda	
		Joao Daniel	R\$ 604.147,05	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	844748	R\$ 494.100,00	Mun. de São Cristóvão	Min. das Cidades
Descrição		Pavimentação de diversas ruas no Município de São Cristóvão/SE.		
Parlamentar		Nome	Valor Emenda	
		Joao Daniel	R\$ 484.891,30	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	845342	R\$ 800.000,00	Mun. de Aquidabã	Min. das Cidades
Descrição		Implantação/melhoria de infra. para mobilidade urbana no mun. de Aquidabã.		
Parlamentar		Nome	Valor Emenda	
		Jony Marcos	R\$ 789.800,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	845722	R\$ 500.000,00	Mun. de Propriá	Min. das Cidades
Descrição		Serviço de obras e pavimentação interligando o bairro propriazinho ao bairro matadouro na sede do município de Propriá.		
Parlamentar		Nome	Valor Emenda	
		Virginio de Carvalho	R\$ 493.100,00	

Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	848513	R\$ 500.000,00	Hospital São José	Min. da Saúde
Descrição		Aquisição de equipamento e material permanente para unidade de atenção especializada em saúde		
Parlamentar		Nome	Valor Emenda	
		Virginio de Carvalho	R\$ 500.000,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	849576	R\$ 99.562,71	Soc. Musical Sra. Santana	Min. da Cultura
Descrição		Aquisição de Instrumentos Musicais, que irão abranger novos alunos do nosso município, da rede municipal, estadual e particular, incentivando-os na aprendizagem da musica de forma a adquirirem valores disciplinares, sociabilidade, convivência em grupos e inclusive social		
Parlamentar		Nome	Valor Emenda	
		Eduardo Amorim	R\$ 99.562,71	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	849588	R\$ 99.587,90	Assoc. Musical Lira Carlos Gomes	Min. da Cultura
Descrição		Aquisição de equip. e instrumentos musicais para atender a demanda do Projeto Escola de Iniciação Musical Lira Carlos Gomes do município de Estância/Se.		
Parlamentar		Nome	Valor Emenda	
		Eduardo Amorim	R\$ 99.587,90	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	852226	R\$ 526.407,18	Estado de Sergipe	Min. do Turismo
Descrição		Obra de reforma da casa de cultura Silvio Romero no Mun. de Lagarto/SE.		
Parlamentar		Nome	Valor Emenda	
		Fabio Reis	R\$ 487.500,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	857874	R\$ 500.800,00	Estado de Sergipe	Min. da Cultura
Descrição		Terceira Etapa dos Serviços e Obras de Execução da Reforma da Catedral Metropolitana de Aracaju, no Estado de Sergipe.		
Parlamentar		Nome	Valor Emenda	
		Antônio Carlos Valadares	R\$ 250.000,00	
		Eduardo Amorim	R\$ 250.000,00	
Ano	Nr. Conv.	Vl. Global	Proponente	Área
2017	859022	R\$ 570.820,00	Estado de Sergipe	Min. Da Cultura
Descrição		Realização do Festival de Ritmos Brasileiros - O Rasgadinho		
Parlamentar		Nome	Valor Emenda	
		Jony Marcos	R\$ 564.820,00	
Favorecido		Nome	Qtd.	Vl. Fatura
		Benguela Prod. e Eventos Culturais LTDA	1	R\$ 28.731,00
		Cactus Promoções E Produções Eireli	1	R\$ 28.500,00
		Estreia Prod. Culturais E Eventos LTDA	1	R\$ 47.500,00
		Garota Sangue Bom Produções Artísticas e Edições Musicais LTDA	1	R\$ 47.542,53
		Jerru Com. e Serv. De Consultoria Emp. LTDA	1	R\$ 27.413,22
		Lado B Prod. Artísticas e Proj. Culturais LTDA	1	R\$ 47.500,00
		Luan Granello Singh Prod. e Eventos	1	R\$ 27.866,66
		Luiz Jose Fontineli de Souza : 590.103.635-20	1	R\$ 8.102,50
		Marya Bunita Produções Artísticas LTDA	14	R\$ 133.697,48
		S. L. Pessoa - Shows	1	R\$ 24.232,50
		Suzana Vasconcelos De Moraes Rego : 335.783.975-04	1	R\$ 2.300,00
		Teo Santana Emp., Propaganda e Eventos LTDA	1	R\$ 7.697,38
		Villela Produções E Emp. Artísticos Eireli	1	R\$ 7.697,38
		W.F.A. Eventos Sociedade Simples LTDA	1	R\$ 49.000,00